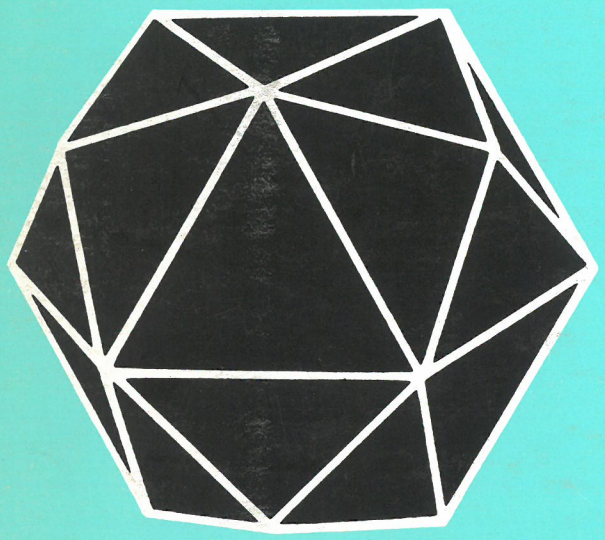
Quaderni di informatica

Rassegna di tecnologia ed applicazioni degli elaboratori elettronici

a cura del Centro Studi Informatica e Automazione Bull Italia



ARCHIVIO STORICO
ASSOCIAZIONE POZZO DI MIELE
FPDM-42
RQDI-133

35
Anno XV - Numero 1 - 1989

FPDM-42

Worldwide
Information
Systems

Bull

Dall'inizio del corrente anno, la Honeywell Bull Italia, editrice di questa rivista, è entrata a far parte del Gruppo Bull, una nuova entità mondiale del settore informatico.
In conseguenza, la società ha assunto il nome di Bull Italia.
Come vi è totale continuità nel nuovo contesto per quanto riguarda le attività scientifiche, industriali e commerciali svolte dall'azienda, così rimane inalterato l'obiettivo della rivista di costituire un veicolo per la diffusione di cultura informatica nel nostro Paese. Un ruolo che "Quaderni di Informatica" svolge ormai da 15 anni, con ampio e lusinghiero riscontro.

Quaderni di informatica

Rassegna di tecnologia ed applicazioni degli elaboratori elettronici
Anno XV - Numero 1 - 1989

Sommario

Dove occorre il supercomputer

In molti settori della scienza e della tecnica esistono problemi la cui soluzione richiede potenze di calcolo di alcuni ordini di grandezza superiori a quelle attualmente disponibili. La risposta a queste esigenze sta nella realizzazione di architetture parallele.
L. DADDA

Perchè il controllo di qualità del software

Conviene investire nella qualità del software? A questa domanda risponde, in modo ampiamente documentato, uno dei maggiori esperti dell'argomento.
M. MAIOCCHI

Microelettronica per la sicurezza: le "smart cards"

L'uso di tessere intelligenti a microprocessore sta diffondendosi rapidamente. Esse offrono, infatti, un livello di sicurezza di gran lunga superiore ai sistemi tradizionali.
A. BRIOSCHI, A. GROSSI

CAI per l'handicap: un supporto alla formazione professionale

Viene descritto un progetto per l'avviamento professionale di portatori di handicap, in cui il computer svolge un ruolo significativo.
M. CANTARELLA

Un anniversario storico: la teoria dell'informazione

Sono passati quarant'anni dalla pubblicazione di un famoso articolo in cui Claude Shannon gettava la basi della teoria dell'informazione.
W. ASPRAY

Direttore Responsabile
Franco Filippazzi

Comitato di Redazione
A. Cicu, C. Falcetti, P. Lupo,
M. Nobile, G. Occhini
F. Sala

Redazione
Bull Italia
Centro Studi Informatica e Automazione
Via G.M. Vida, 11 - 20127 Milano

Stampa
tipolito Maggioni srl

Autorizzazione del Tribunale di Milano
n. 93 del 20 Marzo 1974

«Quaderni di Informatica» ospita articoli e contributi di varia provenienza. La responsabilità delle opinioni espresse rimane agli Autori.

La riproduzione di articoli della rivista, o di parte di essi, è consentita solo citando la fonte e l'autore.

Dove occorre il supercomputer (*)

LUIGI DADDA
Politecnico di Milano

1. Premessa

Una introduzione alla elaborazione parallela può essere fatta ponendosi da diversi punti di vista.

Un tema molto naturale sarebbe una rassegna sulla evoluzione del concetto e delle realizzazioni relative al "supercalcolo", cioè delle teorie, delle macchine e delle applicazioni che hanno, nel corso dei passati decenni, teso a spingere sempre più in là la "potenza" delle macchine informatiche al fine di affrontare problemi prima non risolubili.

È opportuno osservare a tale proposito che nel passato tale processo è stato prevalentemente governato dall'offerta: gli informatici hanno via via concepito e realizzato macchine più veloci, gli utilizzatori le hanno applicate a problemi via via più complessi esercitando una continua spinta sui primi. Spinta che si è manifestata attraverso il classico meccanismo offerta-domanda e che ha prodotto una evoluzione delle macchine quasi continua: memorie e circuiti via via più veloci e capaci, architetture che possono forse meglio definirsi come "incrementi architetturali" piuttosto che "rivoluzioni architetturali": le unità aritmetiche parallele, quelle vettoriali (che non sono altro che strutture SIMD: Single Instruction-Multiple Data), le memorie gerarchiche (cache), una moderata molteplicità di unità di calcolo, ecc. Il tutto nell'ambito della classica architettura sequenziale, alla Von Neumann.

Tutto ciò è confermato dal successo di macchine come quelle prodotte dalla Cray, che costituiscono la stragrande maggioranza dei supercalcolatori oggi installati in tutto il mondo.

Quanto detto sopra non significa peraltro che le ricerche e gli esperimenti su architetture nuove, fortemente innovative rispetto a quelle di Von Neumann, siano mancate. Esse hanno anzi caratterizzato la ricerca informatica lungo tutti i decenni trascorsi, ed è proprio ciò che dovrebbe essere messo in luce in un discorso di prospettiva storica.

Poichè d'altra parte tali tipi di discorso non sono mancati, è sembrato più opportuno esporre alcune considerazioni basate più su una prospettiva futura e ciò in quanto essa è ritenuta più utile per l'orientamento delle ricerche nel campo dei supercalcolatori.

La ragione principale in appoggio a tale punto di vista può essere individuata nella ormai da tempo constatata necessità di uscire dal cosidetto "Von Neumann bottleneck". La limitazione intrinseca alle architetture essenzialmente sequenziali, già da tempo individuata dagli informatici, è oggi constatabile con chiarezza anche dal mondo degli utenti.

Ciò rappresenta una novità, in quanto vede questi ultimi esprimere una domanda non più sostanzialmente indifferenziata e generica verso potenze più alte, ma l'indicazione di traguardi abbastanza bene individuati e di gran lunga superiori a quelli che gli informatici potranno essere in grado di offrire nel futuro prossimo: si tratta in altre parole di una vera e propria sfida che il mondo della ricerca in generale lancia al mondo informatico.

Come primo tema si considererà il *ruolo del supercalcolo* nella ricerca scientifica e tecnologica, anche con lo scopo di individuarne le esigenze.

Un secondo tema da trattare è poi costituito dalla esigenza di sistemi a *forte parallelismo* come strada obbligata per una risposta alle esigenze espresse dalla ricerca scientifica e tecnologica.

Un terzo tema riguarda i problemi di *visualizzazione* dei risultati che si accompagnano all'accresciuta potenza di calcolo e che debbono essere affrontati a scala ben maggiore di quella seguita nel passato.

Analogo discorso può essere fatto a riguardo dei problemi di *comunicazione,* che richiederanno soluzioni adeguate alle potenze delle nuove macchine.

Per quanto riguarda la situazione attuale, la tabella A mostra una classificazione delle macchine oggi in uso.

2. Il ruolo del supercalcolo nella ricerca

Sul versante delle *esigenze*, le previsioni che riguardano il mondo dei

Relazione di apertura del convegno AICA-AEI su "Elaborazione parallela" (Milano, 17-18 novembre 1988).

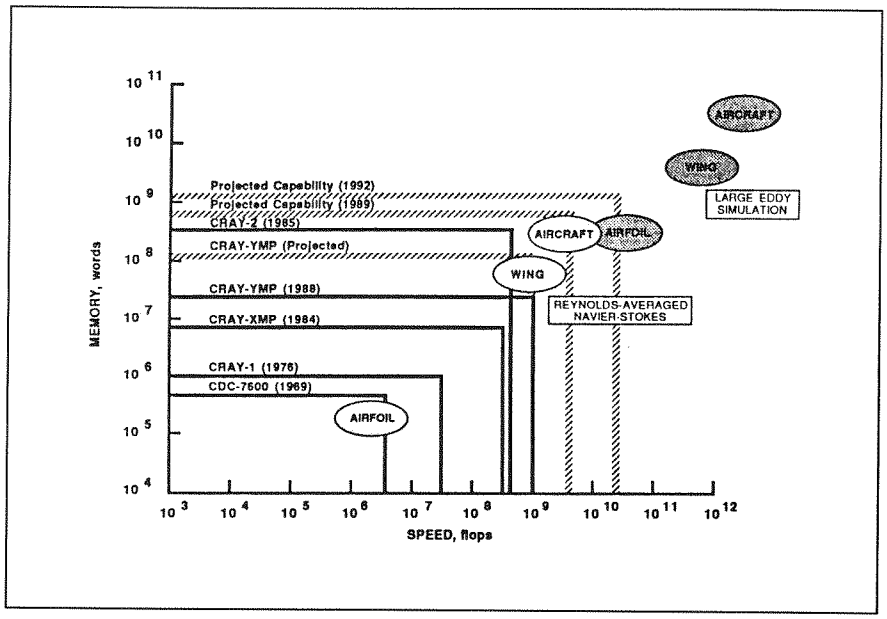


Fig. 1 - Requisiti computazionali (velocità di elaborazione e capacità di memoria) nel settore della AERODINAMICA (Fonte: NASA)

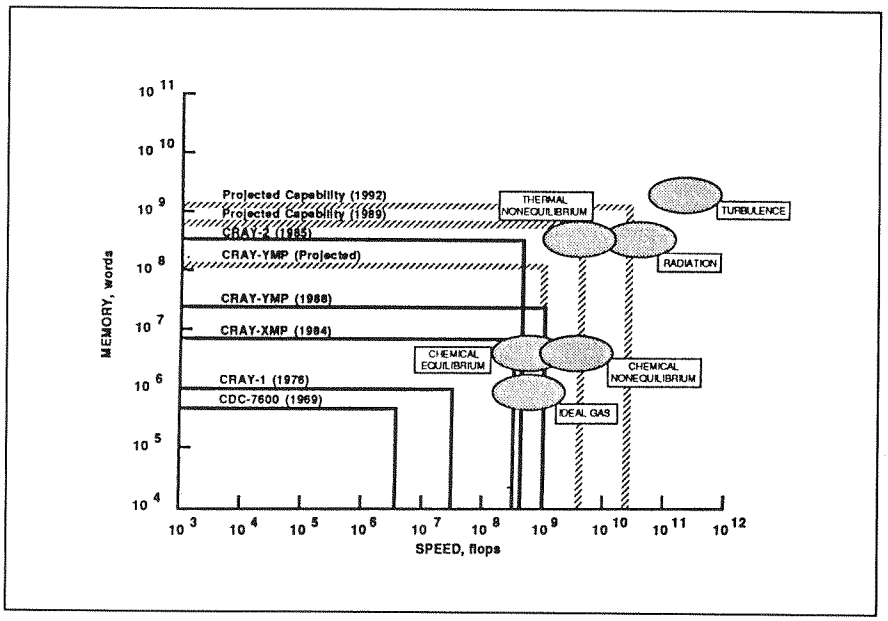


Fig. 2 - Requisiti computazionali nel settore della AEROTERMODINAMICA (Fonte: NASA)

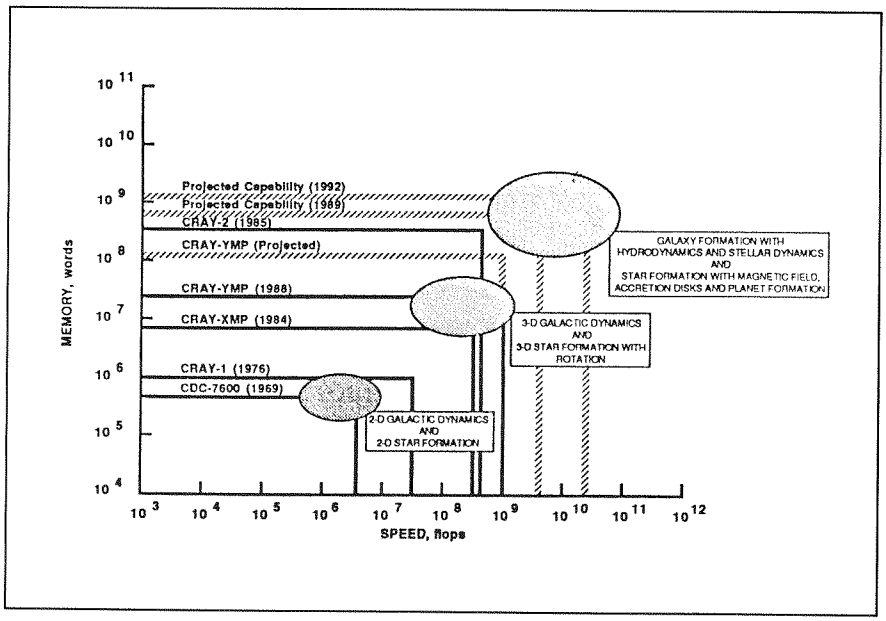


Fig. 3 - Requisiti computazionali nel settore della ASTROFISICA (Fonte: NASA)

supercalcolatori non sono espresse dalla semplice statistica delle opinioni fornite da singoli ricercatori, o da ricerche di mercato che si basino, come è per esse buona norma, sui dati relativi agli anni trascorsi e su proiezioni più o meno prudenziali. L'affermazione della *simulazione numerica* come metodo autonomo per la soluzione di problemi scientifici ed ingegneristici ha non solo fornito dati sperimentali essenziali per la valutazione delle esigenze, ma ha indotto a veri e propri studi di proiezione a medio-lungo termine nei più svariati settori scientifici e tecnologici. Tali studi hanno messo in evidenza problemi la cui soluzione non può essere ottenuta con altri mezzi che non siano quelli numerici. Il fisico Wilson [3] chiama tali problemi le "grandi sfide" (grand challenges) della scienza, e pensa che almeno due di esse, quella riguardante la struttura elettronica e la turbolenza potrebbero avere un impatto paragonabile a quello della scoperta di nuove leggi della natura.

Altri "grandi sfide" sono state individuate nelle aree della biologia molecolare, dell'aerodinamica, della teoria quantistica dei campi, dei sistemi economici.

Sono comunque tutti problemi non risolubili con gli odierni supercalcolatori. Basta osservare le figure da 1 a 12, che riportano i risultati degli studi di previsione nei seguenti campi:

1-aerodinamica,

2-aerotermodinamica,

3-astrofisica,

4-scienza dell'atmosfera,

5-chimica,

6-impatti ad ipervelocità,

7-grandi (future) strutture spaziali,

8-ottica laser,

9-magnetoidrodinamica,

10-visualizzazione scientifica (sulla quale più avanti si ritornerà per l'importanza da essa rivestita per quanto riguarda la tecnica del supercalcolo),

11-simulazione del sistema strategico di difesa,

12-ricerche sulla visione.

Vi si nota come le velocità di calcolo e la capacità di memoria richiesta debba crescere di alcuni ordini di grandezza rispetto ai valori odierni. L'obiettivo di supercalcolatori capaci di un *teraFLOPS* (mille miliardi di FLOPS) si pone perciò come l'obiettivo da raggiungere per poter vincere le grandi sfide.

I supercalcolatori della attuale generazione realizzano all'incirca un miliardo di FLOPS, e comprendono al più 8 processori. Quelli progettati per il 1992 ne comprendono 64. È facile concludere che l'obiettivo anzidetto non potrà essere raggiunto con le attuali architetture. È pertanto necessario pensare a macchine con un parallelismo molto più spinto, un parallelismo "massiccio": è questa un'altra indicazione che emerge da quanto sopra detto.

Le "grandi sfide" scientifiche si traducono così in grandi sfide tecnologiche.

Quanto è stato finora detto basta a confermare la convinzione che noi siamo oggi non tanto nella scia di una rivoluzione informatica iniziata da oltre un quarantennio e che prosegue, bensí nel bel mezzo di una "rivoluzione computazionale" che sta cambiando la scienza e di conseguenza la società in modo altrettanto profondo a quello operato dalle rivoluzioni agricola ed industriale. Questi sono più o meno i termini con cui si esprimono numerosi ricercatori attivi nel campo delle scienze computazionali. Si può non condividere l'estremismo di tali affermazioni, ma occorre comunque riconoscere che il modo di fare scienza e tecnologia sta cambiando con l'avvento del supercalcolo.

Vi è chi, come prima reazione, avverte gravi pericoli nella pratica della simulazione come vero e proprio attentato al metodo scientifico galileiano ed ai suoi fondamenti di sperimentazione come interrogazione della natura e di formulazione di teorie come linguaggio di interpretazione della stessa, e parlano di "sindrome da macchina incantata".

È peraltro facile convincersi che il calcolo era certamente secondario o addirittura inesistente ai primordi del metodo scientifico, mentre esso è diventato via via sempre più utile fino a diventare, oggi, semplicemente indispensabile. Per verificare la costanza dell'accelerazione nella caduta dei gravi bastò un esperimento ingegnoso e semplice, ed il relativo modello matematico non richiedeva nessun calcolo. Leggere nelle odierne teorie gli esperimenti atti a legittimarle, nonché governare gli esperimenti relativi, è diventato compito semplicemente impossibile senza il ricorso al supercalcolo. È ciò che spiega la nascita, in questi anni, delle varie "scienze computazionali".

Il calcolo si affianca così alle formulazioni teoriche ed agli apparati sperimentali come il terzo pilastro del metodo scientifico.

Tutto ciò merita certamente una più accurata discussione critica, dal momento che si affronta un problema epistemologico di base. Si può osservare in aggiunta a quanto già detto, che l'uso del calcolatore come "abaco veloce", caratteristico dei trascorsi decenni, non ha contribuito che indirettamente allo sviluppo delle conoscenze scientifiche e delle applicazioni tecnologiche. Non ci si aspettava dal calcolo elettronico contributi concettuali, ma soltanto la rimozione degli ostacoli costituiti dalla lunghezza dei calcoli, impossibili da svolgere coi mezzi precedenti. Con altre parole, il significato dei risultati dei calcoli era perfettamente noto dalla teoria che li aveva indicati.

Tutto ciò ha avuto un'importanza grandissima sul modo di condurre le ricerche, ed anche per rivitalizzare settori scientifici, spesso bloccati nel loro sviluppo dal loro grado di complessità, impossibile da affrontare con i mezzi della pura analisi matematica. Un esempio è dato dai sistemi dinamici non lineari.

Sono nel frattempo emerse nuove ed ancora più complesse esigenze di "matematizzazione" in scienze come la fisica stessa, la chimica, la biologia, l'economia. È stata quasi una regola nel passato il trovare strumenti matematici già pronti nel momento in cui le scienze fisiche si imbattevano in problemi nuovi ed inconsueti. Si può dire che la matema-

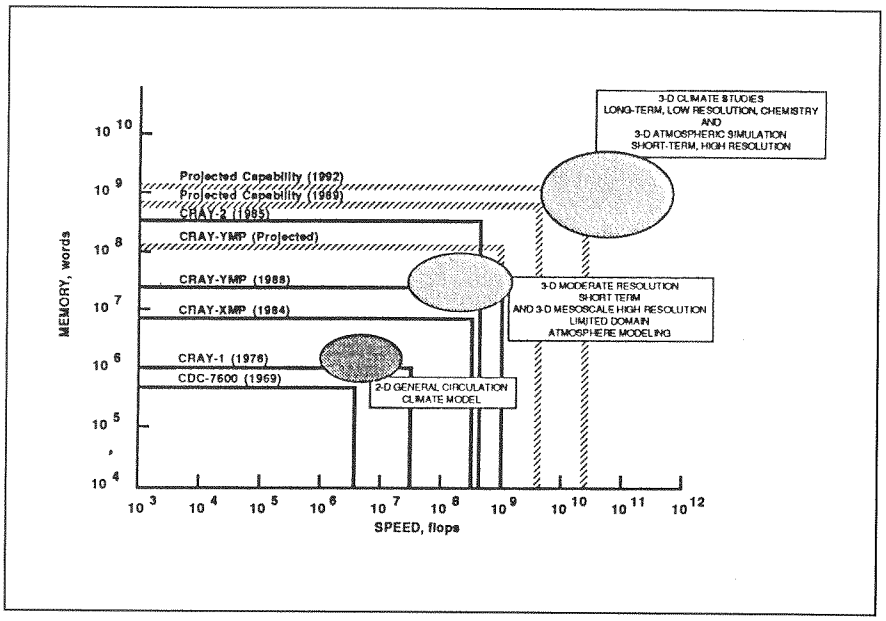


Fig. 4 - Requisiti computazionali nel settore delle SCIENZE DELL'ATMOSFERA (Fonte: NASA)

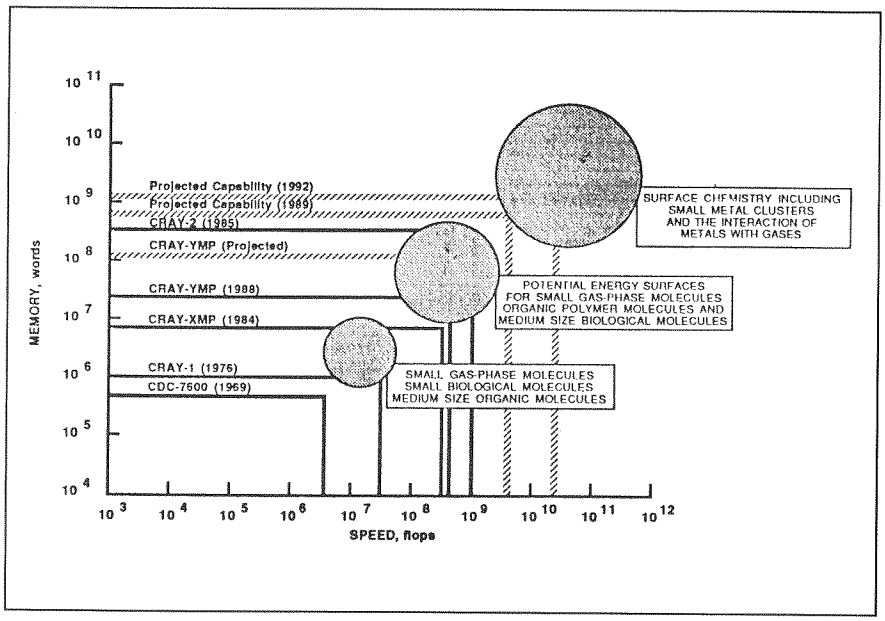


Fig. 5 - Requisiti computazionali nel settore delle CHIMICA (Fonte: S. Diego Supercomputer Center)

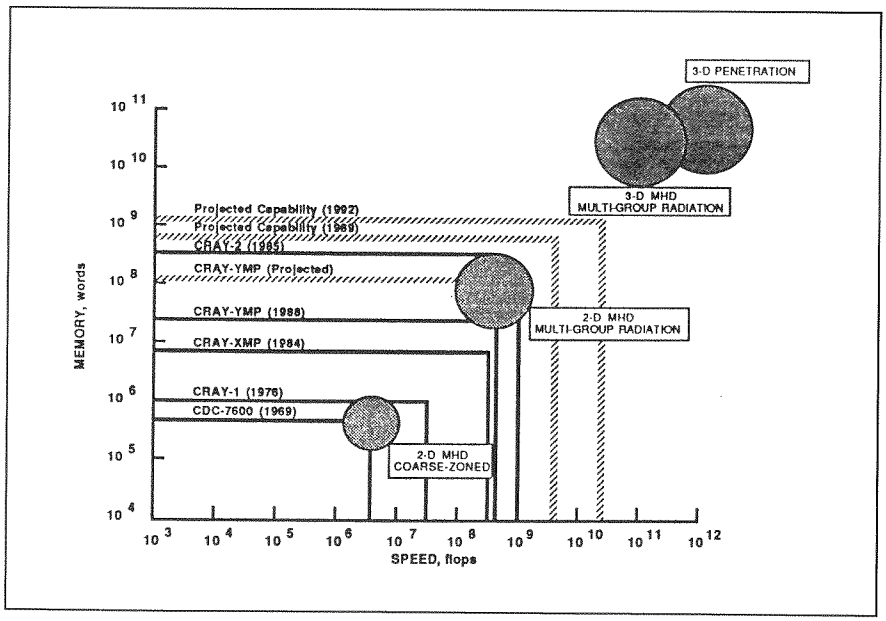


Fig. 6 - Requisiti computazionali nel settore degli IMPATTI ad IPERVELOCITÀ (Fonte: USAF)

tica fu spesso in grado di battere in anticipo i sentieri di sviluppo delle discipline fisiche ed applicative. Forse ciò è oggi un po' meno vero.

Del resto, la stessa matematica ha da confrontarsi più "direttamente" col calcolatore: non tanto cioè per le sue capacità puramente numeriche, ma come strumento "logico". Anche se i primi esempi lo fanno pensare ancora come un "abaco veloce": basta pensare al classico problema dei quattro colori, la cui dimostrazione è stata ottenuta (più di dieci anni fa) tramite un calcolatore, usato per la verifica delle condizioni, in numero molto grande, alle quali la dimostrazione era stata condotta. La dimostrazione automatica di teoremi è, peraltro, uno dei temi fondamentali nell'ambito delle ricerche di intelligenza artificiale.

È comunque nelle scienze sperimentali e nelle ingegnerie che il calcolatore sta sempre più intensamente intervenendo con la *simulazione* dei fenomeni naturali nelle prime e dei sistemi artificiali di grande complessità nelle seconde. La simulazione, come già detto, si affianca alla teoria ed all'esperimento come strumento di indagine della natura. Il ruolo della teoria nella individuazione di nuovi fenomeni da confermare con l'esperimento viene esteso alla simulazione, dalla quale possono emergere aspetti nuovi ed inattesi, non facilmente deducibili dal modello teorico, da sottoporre alla verifica dell'esperimento. Inoltre, possono verificarsi casi nei quali lo stesso esperimento è difficile da attuarsi o da interpretare e la simulazione rappresenta la sola operazione realizzabile in appoggio alla stessa teoria.

Ciò rappresenta una delicata, e perciò controversa, fase di passaggio da una tradizione che vedeva nelle equazioni fondamentali di un fenomeno la forma perfetta della conoscenza del mondo fisico ad un'altra nella quale un fenomeno è considerato noto solo quando ne siano stati conosciuti i meccanismi di base (contenuti nelle equazioni, ma in forma non evidente). Ciò può essere ottenuto con la simulazione di un modello (invece che scrivendo le equazioni dello stesso), formulando così un esperimento ideale (*gedankexperiment*). L'obiezione che si può rivolgere a tale metodo è che così facendo si studia una realtà artificiale, esistente solo nella mente del ricercatore e da esso riprodotta nel calcolatore. Ma ciò è vero anche per un sistema di equazioni, essendo la verifica sperimentale nell'un caso e nell'altro la sola legittimazione accettabile.

La ricerca nelle scienze naturali è oggi rivolta con intensità crescente ai sistemi caratterizzati da grande complessità, come i sistemi biologici (per es. le reti di neuroni), la fisica dello stato solido (i difetti nei cristalli, i materiali amorfi quali il vetro), per i quali non esistono strumenti matematici adeguati. L'atteggiamento tradizionale di fronte a tale genere di problemi è stato di ritenerli "poco nobili" in quanto non descrivibili con teorie pulite e pertanto da relegare nel campo delle meno nobili applicazioni, nelle quali non ci sarebbe nulla da capire dal punto di vista strettamente scientifico.

Non meno rilevanti, anche se di diversa natura, sono le implicazioni ingegneristiche del supercalcolo, nelle quali si ha insieme a che fare appunto con i "meno nobili" problemi di conoscenza fisica nonché (e questo è l'aspetto più caratteristico) con la concezione di sistemi artificiali di enorme complessità. Tra i primi si possono annoverare i problemi di fluidodinamica ed in particolare i moti vorticosi, nei secondi le macchine e gli impianti di crescente complessità che caratterizzano l'ingegneria moderna in tutte le sue branche: dalle costruzioni alle stesse macchine informatiche, non dimenticando gli aspetti che se anche non di dominio della ingegneria come tradizionalmente intesa sono correttamente aggettivate come ingegneristiche in quanto ispirate dalla stessa "forma mentis", quale l'ingegneria delle molecole complesse e la stessa ingegneria genetica.

La stessa scienza delle macchine, una volta essenzialmente tributarie di una sola disciplina (meccanica, elettrotecnica, ecc.), è sempre più pluridisciplinare, essendovi inclusa anche la nuova disciplina informatica che più di ogni altra caratterizza le nuove macchine conferendo ad esse "intelligenza".

3. Le macchine a forte parallelismo. Gli algoritmi paralleli

È ben noto che i calcolatori tradizionali siano stati concepiti come "macchine algoritmiche", fin dall'epoca di Babbage. Turing, anzi, diede degli algoritmi una definizione basata proprio sullo schema di una macchina ideale (di concezione molto elementare, e comunque di tipo sequenziale), essendo l'algoritmo definito come il programma della stessa. Ciò è del resto diventato costume nelle comunicazioni di risultati tra ricercatori, già da alcuni decenni, con l'uso di linguaggi standard, come il Fortran.

L'algoritmo è stato così sempre pensato e descritto in termini rigorosamente sequenziali. Ciò spiega, tra l'altro, l'immediato successo goduto dal calcolatore nel campo scientifico. Si può dire, con altre parole, che il calcolatore tradizionale ha trovato nella ricerca matematica e fisica dei secoli trascorsi un immenso patrimonio di algoritmi che attendevano solo lo strumento adatto per dispiegare i loro effetti nelle applicazioni e nella ricerca stessa.

La strada delle macchine parallele, anzi a forte parallelismo, obbliga a riesaminare l'impostazione generale data sia alla concezione delle macchine sia agli stessi algoritmi.
È stato giocoforza sviluppare i supercalcolatori secondo l'architettura tradizionale, anche perché gli stessi algoritmi sono stati su di essa definiti e, si può dire, plasmati. Ma è altrettanto giocoforza oggi pensare ad architetture ad alto parallelismo, le uniche che potranno permettere di vincere le nuove, grandi sfide. Su di esse sappiamo ancora poco e, soprattutto, sappiamo ancora meno sul modo di concepire gli algoritmi per esse adatti. Si è anzi capito che si impone un approccio interdisciplinare che veda affiancati nella ricerca gli architetti informatici nelle varie accezioni (esperti di tecnologie, di hardware, di software di base, di linguaggi), ai matematici e agli esperti degli specifici settori applicativi in-

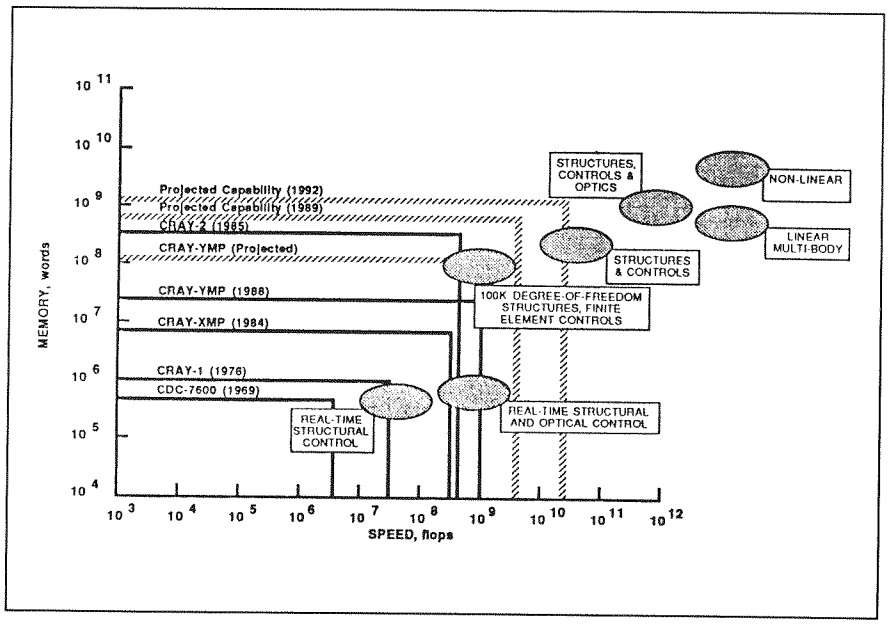


Fig. 7 - Requisiti computazionali nel settore delle GRANDI STRUTTURE SPAZIALI (Fonte: USAF)

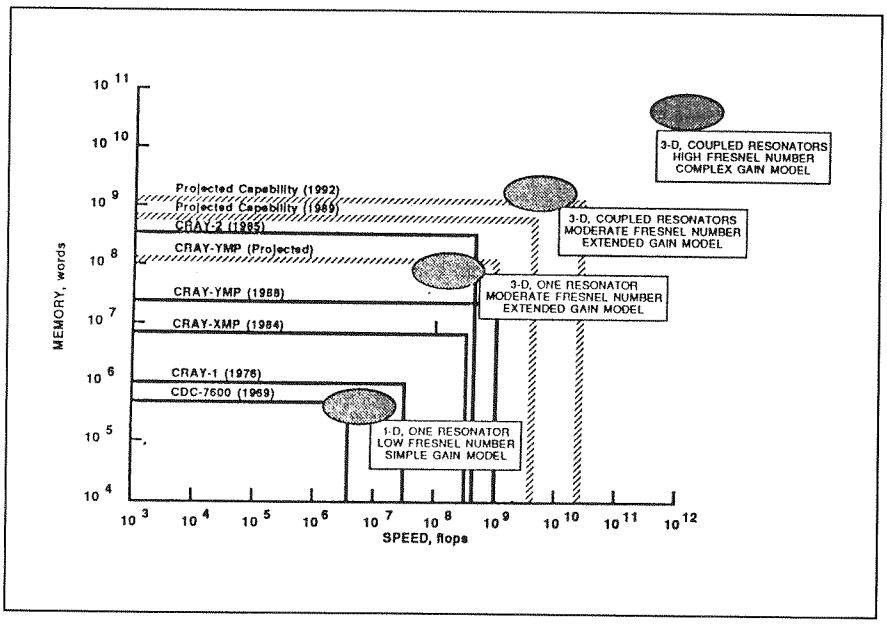


Fig. 8 - Requisiti computazionali nel settore dell'OTTICA LASER (Fonte: USAF)

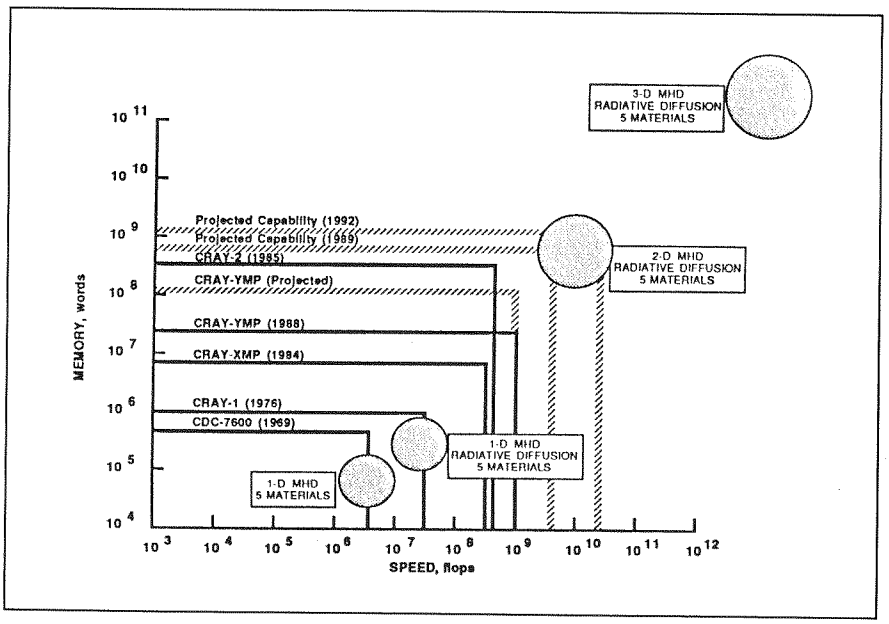


Fig. 9 - Requisiti computazionali nel settore della MAGNETOIDRODINAMICA (Fonte: USAF)

teressati (dalla fluidodinamica, alla chimica, alla biochimica, all'econometria, ecc.).

La fig. 13 mette a confronto le velocità di calcolo effettive misurate su macchine di architettura tradizionale (CRAY, IBM) con quelle di macchine di nuova architettura, a medio ed alto parallelismo. Si tratta di un confronto "stile scienze naturali", su una popolazione abbastanza consistente di macchine, senza ricorso a simulazioni o modelli e su programmi relativi a problemi concreti. Il risultato che se ne può ricavare è che macchine anche ad elevato parallelismo (fino a mille unità) riescono a sfruttare abbastanza bene la potenza complessiva, grazie sia alle architetture adottate sia a nuovi algoritmi studiati in vista del parallelismo sia delle macchine che dei problemi stessi.

Si può qui osservare in generale che il mondo fisico è caratterizzato da importanti aspetti di parallelismo che possono essere rappresentati in nuovi algoritmi. Tali aspetti possono essere distinti in "spaziali" e "funzionali", e permettono la partizione di un problema in sotto-problemi eseguibili in parallelo.

Il parallelismo spaziale si ha quando gli effetti di una perturbazione si fanno sentire soprattutto nelle immediate vicinanze (si noti che ciò vale anche per la dimensione "tempo"). Si possono allora definire domini spaziali caratterizzati da deboli interazioni, rappresentando ciascuno di essi con un processore cui si richiede lo scambio di informazioni solo con quelli vicini. Gli esempi applicativi sono numerosi: basta pensare a tutti i problemi che possono modellarsi con i metodi degli elementi finiti, così importanti anche in campo ingegneristico.

Il parallelismo funzionale si ha invece quando i fenomeni fisici rappresentati dal modello considerato sono caratterizzati da scale spaziali o temporali alquanto diverse. Si può allora pensare di assegnare tali fenomeni quasi-distinti a diversi processori (o gruppi di processori) tra i quali è previsto un relativamente debole scambio di informazioni. Esempi di parallelismo funzionale si trovano in sistemi fluidodinamici nei quali gli effetti della viscosità vengono supposti limitati agli strati limite, nonché in reazioni chimiche che avvengono in flussi a più componenti.

Lo sviluppo di algoritmi intrinsecamente paralleli è uno dei compiti più importanti ed ardui del prossimo futuro, che richiederà tra l'altro il superamento dell'attuale, diffusa cultura del "pensare scalare".

Allo sviluppo degli algoritmi paralleli deve accompagnarsi ovviamente quello dei linguaggi e degli strumenti di programmazione. L'esperienza in corso ha visto le estensioni di linguaggi tradizionali, quali il Fortran, ma si possono prevedere sistemi operanti ad un più alto livello di astrazione, probabilmente con l'adozione di sistemi esperti.

4. Il problema della "visualizzazione" dei risultati

L'immagine del programmatore che ritorna dal centro di calcolo carico di carta stampata coi risultati di un programma è viva nella memoria di molti di noi, anche perché è sempre stata accompagnata dal dubbio sulla utilità di estrarre dalla macchina tutti quei dati, per l'assoluta impossibilità di poterli esaminare (se non scorrendo le colonne di numeri per scoprire quando appariva un segno negativo, od una cifra non nulla). Essa è stata poi sostituita da quella dello stesso programmatore seduto davanti ad un terminale video, dove i dati possono essere agevolmente esaminati, scegliendo quelli da stampare. L'avvento dei supercalcolatori ripropone tuttavia, ed in forma più grave, lo stesso problema: come esaminare in modo fruttifero i risultati di un calcolo, che in un supercalcolatore possono essere espressi anche da molti miliardi di cifre. Basta pensare a problemi di tipo fluidodinamico. È facile giungere alla conclusione che è perfettamente inutile stampare numeri, dai quali nessun essere umano saprebbe estrarre le informazioni veramente importanti, e che non si può fare altro che mettere in gioco il senso umano dotato della maggior "banda passante", che è la vista. Ecco perché si parla di "visualizzazione dei risultati" come importante problema da risolvere per poter efficacemente usare un supercalcolatore.

Per lo stesso motivo diventa rilevante il problema della memorizzazione dei risultati, che riguarda le memorie di massa, che dovranno essere di capacità adeguate, molto maggiori di quelle finora in uso comune, come i dischi magnetici.

L'importanza del problema della visualizzazione emerge bene dal rapporto "Visualization in Scientific Computing" pubblicato recentemente in USA dalla National Science Foundation. Si valuta che ben il 95% dei calcoli compiuti nei centri dotati di supercalcolatori può considerarsi perso per l'impossibilità di interpretare i dati prodotti "esplosivamente" dalle nuove macchine. Non si tratta soltanto di rappresentare i risultati con grafici di vario tipo, resi ben noti anche attraverso i personal computers, ma di escogitare nuove macchine e nuovi modi di rappresentazione visiva che permettano al ricercatore di cogliere rapidamente l'andamento del fenomeno simulato (perché di tale genere di problemi in genere si tratta) allo scopo di permettergli di intervenire per pilotare la simulazione stessa nella maniera più significativa. Il problema richiede un approccio interdisciplinare, in quanto la sua soluzione mette in gioco non solo le competenze specifiche di carattere informatico e quelle del settore scientifico cui il programma si indirizza, ma anche competenze psicologiche e fisiologiche, e, perché no?, artistiche.

Si noti poi che il problema della "esplosione dei dati" non riguarda soltanto i supercalcolatori, ma anche i sempre più potenti sensori che producono i "dati primi", come le sonde spaziali ed in generale i metodi di rilevazione e studio di fenomeni attraverso immagini.

5. I supercalcolatori e le comunicazioni

I supercalcolatori sono per definizione le macchine più potenti e per-

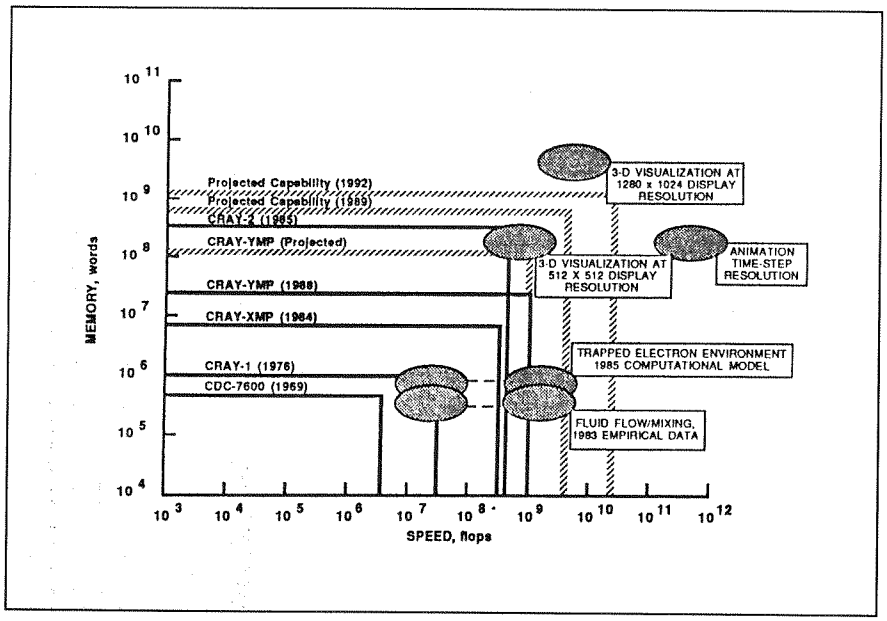


Fig. 10 - Requisiti computazionali nel settore della VISUALIZZAZIONE SCIENTIFICA (Fonte: USAF)

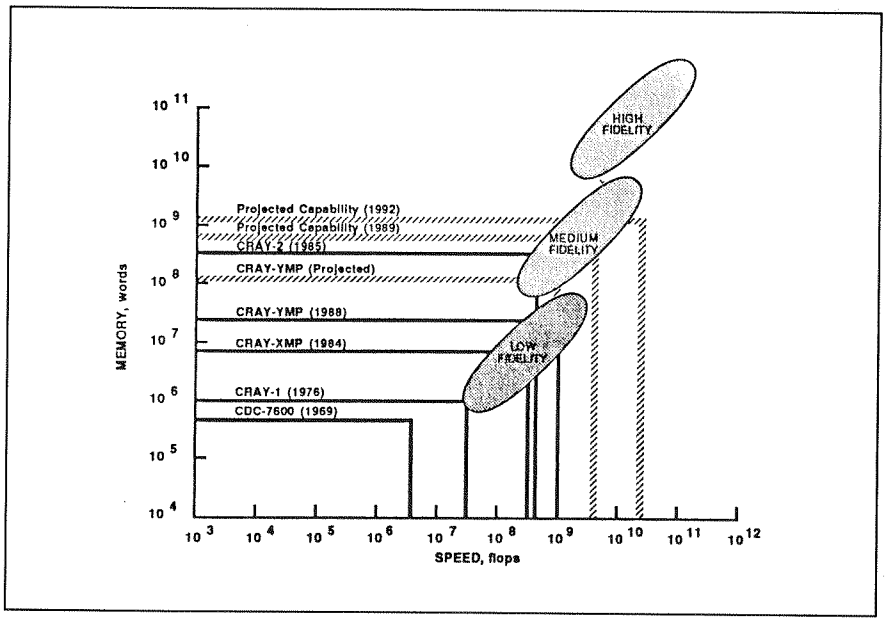


Fig. 11 - Requisiti computazionali nel settore della simulazione del SISTEMA STRATEGICO DI DIFESA (Fonte: Cray Research, Inc.)

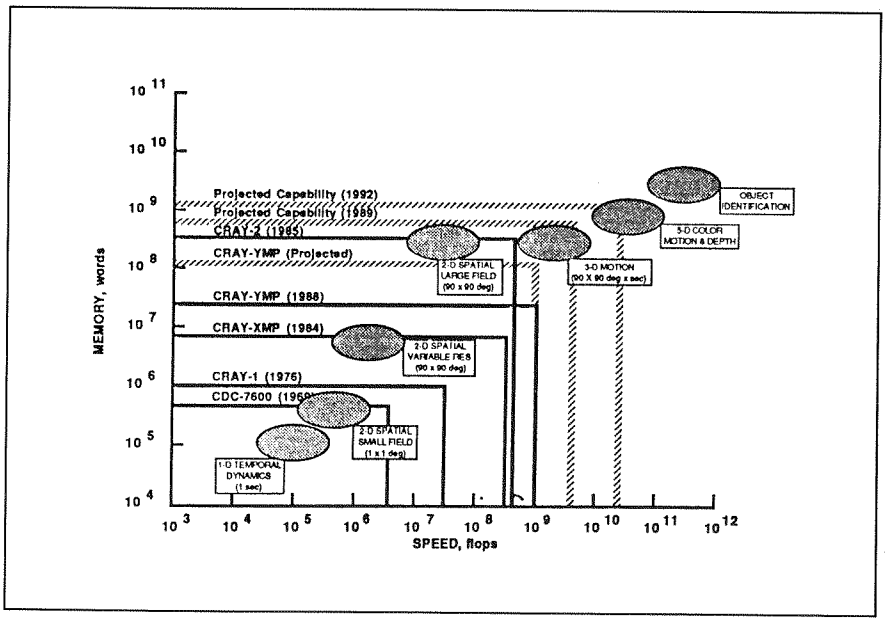


Fig. 12 - Requisiti computazionali nel settore delle ricerche sulla VISIONE (Fonte: NASA)

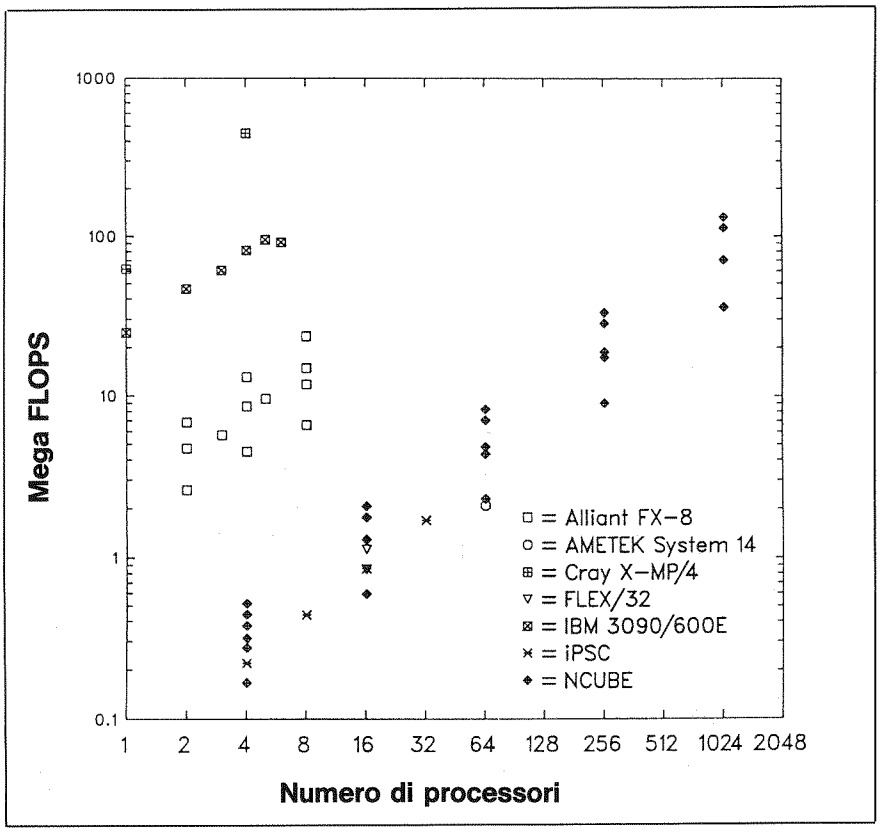


Fig. 13 - Confronto tra le velocità di calcolo misurate su macchine di architettura tradizionale (CRAY, IBM) e quelle di macchine di nuova architettura, a medio ed alto parallelismo.

ciò costose che si possono costruire. Si tratta pertanto di una risorsa "rara", necessaria per lo sviluppo della ricerca in un numero crescente di campi, come è stato detto anche precedentemente.

L'uso a distanza tramite terminali dovrà perciò essere possibile come lo è stato nel passato. Ciò richiederà una rete di trasmissione di caratteristiche adeguate.

Una stima della *banda passante* è stata compiuta [6] sulla base del volume e del tipo di traffico presumibilmente richiesti da un centro di calcolo munito di supercalcolatore, con estrapolazione al prossimo decennio. La stima è stata eseguita con tre modelli, caratterizzati dal tipo di lavori e di utenti. In uno dei modelli, si è supposta l'esistenza di almeno un ricercatore che richieda di ricevere immagini a colori in movimento ad alta risoluzione: una immagine richiede circa 30 megabit di informazioni che, alla frequenza di 30 immagini al secondo, richiedono quasi in gigabit al secondo, 1Gb/s.

In un altro modello si è considerata un'università con 35.000 studenti e 3.000 docenti-ricercatori, pervenendo alla stima di 1Gb/s.

Nel terzo modello si è considerato il caso di un centro che acceda ad apparecchiature remote come supercalcolatori, acceleratori di particelle, laboratori facenti uso di apparecchi per rilievo di immagini anche in movimento, nonché centri per la produzione di immagini in movimento.

TABELLA A: CLASSIFICAZIONE DEI SUPERCALCOLATORI

	sequenziali	paralleli
Super > 100 MFLOPS	Cray X-MP Amdhal/Fujitsu NEC SCX ETA 10 IBM 3090	Intel ipsc-VX N Cube/10 FPS T-series
Mini-super 10-100 MFLOPS	Convex CI SCS-40 FPS M/64 series Cydra	Alliant FX-8 Multiflow
Super-mini < 10 MFLOPS	VAX 8800 IBM 4300 Gould Elxsi	Sequent B21000 Encore Multimax Flexible Flex-32

La larghezza di banda necessaria è risultata ancora di circa 1Gb/s.

Considerata la pratica coincidenza delle stime, si può così assumere una banda minima di 1Gb/s, da aumentare a 3 Gb/s con proiezione a 15 anni.

Si tenga presente a tale proposito che la rete ISDN (Integrated Systems Digital Network) è prevista per servizi richiedenti al più alcuni Mb/s. Di qui la proposta avanzata negli USA per una rete sperimentale che soddisfi le esigenze del supercalcolo.

Bibliografia

[1] *Supercomputers: Government Plans and Policies.* Congress of the United States, Office of Technology Assesment, marzo 1986.

[2] LIVI R., PARISI S., RUFFO S., VULPIONI A., *Il computer: da abaco veloce a strumento concettuale,* Dip.to Fisica Univ. Firenze, 1986.

[3] WILSON, K.G., *Grand Challenges to Computational Sciences,* Cornell University, maggio 1987.

[4] *A National Computing Initiative - The Agenda for Leadership,* SIAM-Soc. for Industr. and Applied Mathematics - National Science Foundation - U.S. Department of Energy, 1987.

[5] Physics today, special issue: *Computational Physics,* ottobre 1987.

[6] *A Research and Development Strategy for High Performance Computing,* Executive Office of the President, Office of Science and Technology Policy, novembre 20, 1987.

[7] JOHNSON, G.M., *Requirements for Parallel Processing in Scientific Computation,* Proc. Third Intern. Conf. Supercomputing, Boston, Mass., maggio 1988.

Perchè il controllo di qualità del software (*)

MARCO MAIOCCHI
Dipartim. di Scienze della Informazione
Università di Milano

1. Il controllo di qualità del software: una definizione

È molto diffusa la visione del testing dei programmi come un'attività dedicata al reperimento degli errori presenti negli stessi [5], e ciò comporta un certo insieme di conseguenze, spesso negative:

- innanzitutto una tale visione privilegia aspetti tecnici e strumentali, che rendano più efficace la produzione di numerosi casi di prova e conseguentemente aumentino la probabilità di trovare errori; anche se ciò è fondamentalmente un aspetto che aumenta la qualità dei programmi, in esso si perde completamente di vista il fatto che un programma è parte di un sistema molto più ampio, e che quello che deve essere garantito è una qualità più globale, che passa anche attraverso la manutenibilità del prodotto completo;
- inoltre tale modo di vedere polarizza il progettista devolto all'attività di testing verso una sorta di "produttività negativa" (il trovare problemi) che lo rende antagonista del progettista che ha sviluppato; ciò è rinforzato dell'atteggiamento del management che ne misura la capacità su tale sorta di produttività inversa, e ciò è anche fonte di una erronea cattiva fama dell'attività di testing come di basso livello e demotivante; di più, l'atteggiamento di antagonismo tra sviluppatore e controllore riduce enormemente l'efficacia delle azioni di testing, sia perchè il secondo tende a nascondere le informazioni al primo, sia perchè questo svolge una limitata azione di misura finale, incapace di influire durante la fase di sviluppo, che è l'unica in cui la qualità può essere creata.

Sposteremo tale ottica considerando l'attività di Controllo di Qualità del Software come quella fase di sviluppo devolta alla *preparazione di strumenti, programmi e dati orientati alla dimostrazione del corretto funzionamento dei programmi e all'acquisizione di un'adeguata confidenza in tale corretto funzionamento* [3].

Con tale ottica l'atteggiamento verso questa disciplina è spostato verso una completa integrazione della stessa con tutte le altre fasi di sviluppo del software, con gli strumenti di sviluppo e con i progettisti.

Riprendiamo alcune definizioni del ciclo di vita del software, come esso risulta schematizzato in fig. 1.

Il processo di produzione del software è un'attività ciclica che parte da

* un certo insieme di *esigenze;* tali esigenze possono essere il risultato di una analisi di mercato, della richiesta di un cliente che richiede di comperare un prodotto, della richiesta di un ente interno a un'organizzazione che, pur in modo non formalizzato da un contratto, diviene cliente;
* attraverso un'**analisi dei requisiti,** produce un documento di *specifiche di requisiti* che individua con precisione i vincoli del problema da risolvere;
* quindi, mediante un'**analisi funzionale,** definisce con dettaglio le *specifiche funzionali,* ovvero le modalità con cui il prodotto interagirà con l'ambiente e con l'utente finale; in tale documento sarà presente anche un'architettura di massima, possibilmente riproducente la struttura dell'architettura funzionale, guida per un'adeguata pianificazione;
* un'attività di **disegno** produrrà quindi un'*architettura di dettaglio,* che sarà il punto di partenza per
* la **codifica,** risultato della quale sarà il *prodotto finale,* che dovrà poi essere sottoposto a
* **qualificazione,** per avere una specifica misura della qualità ottenuta, registrata in un *rapporto di qualità.*

In seguito al *rilascio* del prodotto, inizia un secondo tipo di attività, che vede coinvolte fasi di *vendita,* di *configurazione,* di *personalizzazione,* di *installazione,* di *assistenza* e che infine, tende a evidenziare nuove esigenze (ciò spesso accade per il fatto che la presenza di automazione modifica l'ambiente al punto tale da rendere inadeguata se stessa), che riinnescano il processo di sviluppo.

Rispetto a tale visione, l'attività di controllo di qualità dovrà:

(*) *Il materiale di questo articolo è tratto dal cap. I del volume* «Il controllo di qualità del software» *dello stesso Autore (Ed. F. Angeli, Collana dei Quaderni di Informatica, 1988).*

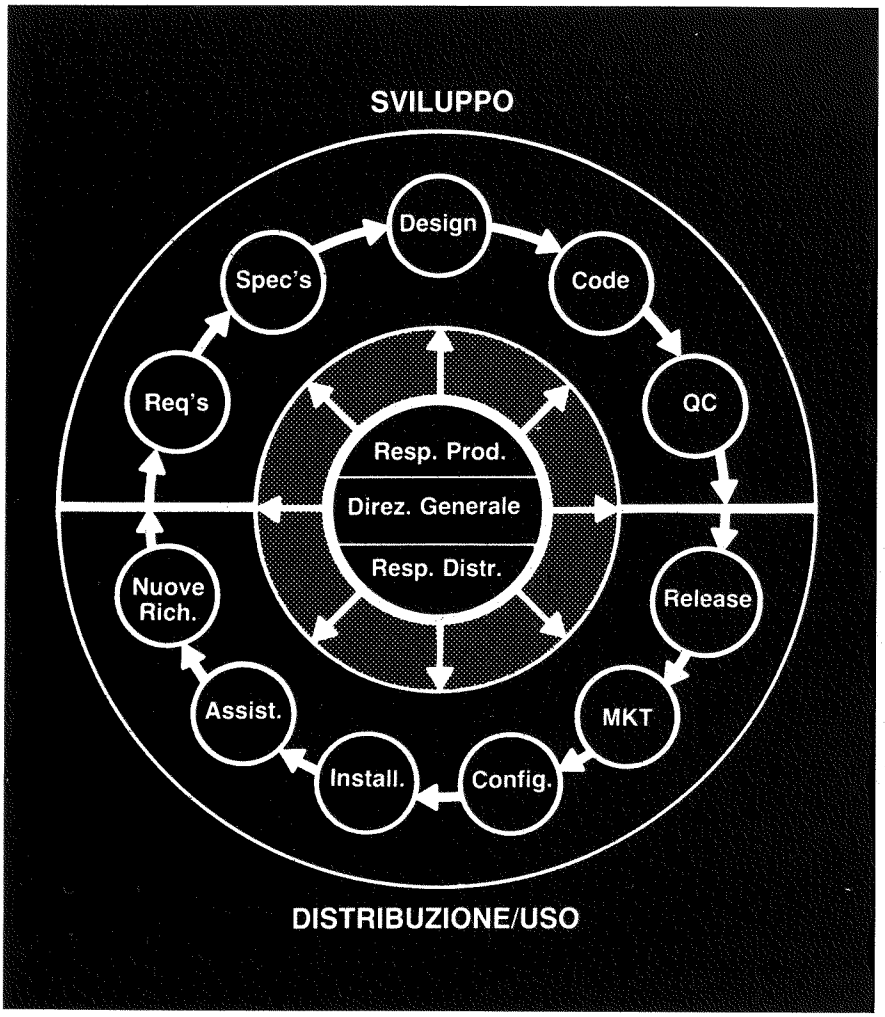


Fig. 1 - Il ciclo di vita del software.

- essere integrata nell'attività di sviluppo con organizzazione, metodi, strumenti;
- garantire la qualità del prodotto così come desiderato dall'utente finale, ovvero dall'organizzazione di distribuzione;
- garantire la qualità del prodotto come desiderato dall'ambiente di sviluppo stesso, in termini di manutenibilità, correggibilità, leggibilità, eccetera.

2. Costi e benefici del QC SW

Da quanto detto emerge come il responsabile di produzione abbia l'obiettivo, a parità di qualità di un prodotto, di ridurre i costi di sviluppo e manutenzione, mentre il responsabile della distribuzione abbia quello di ridurre i costi di gestione, di installazione e di assistenza. Nonostante l'evidente ruolo chiave che hanno nel raggiungimento di tali obiettivi la corretta analisi iniziale e le caratteristiche delle specifiche funzionali del prodotto, esistono alcuni elementi strettamente legati alla morfologia della produzione su cui è possibile influire.

In fig. 2 sono illustrati due tipici diagrammi ampiamente diffusi e noti; dal primo emerge come, nello sviluppo del software l'attività legata al controllo di qualità arrivi a coprire circa la metà dei costi; il secondo mostra come il costo di sviluppo sia solo una parte limitata dei costi totali che dovranno essere previsti per tenere in vita il prodotto, per garantire cioè ciò che viene detto **manutenzione.**

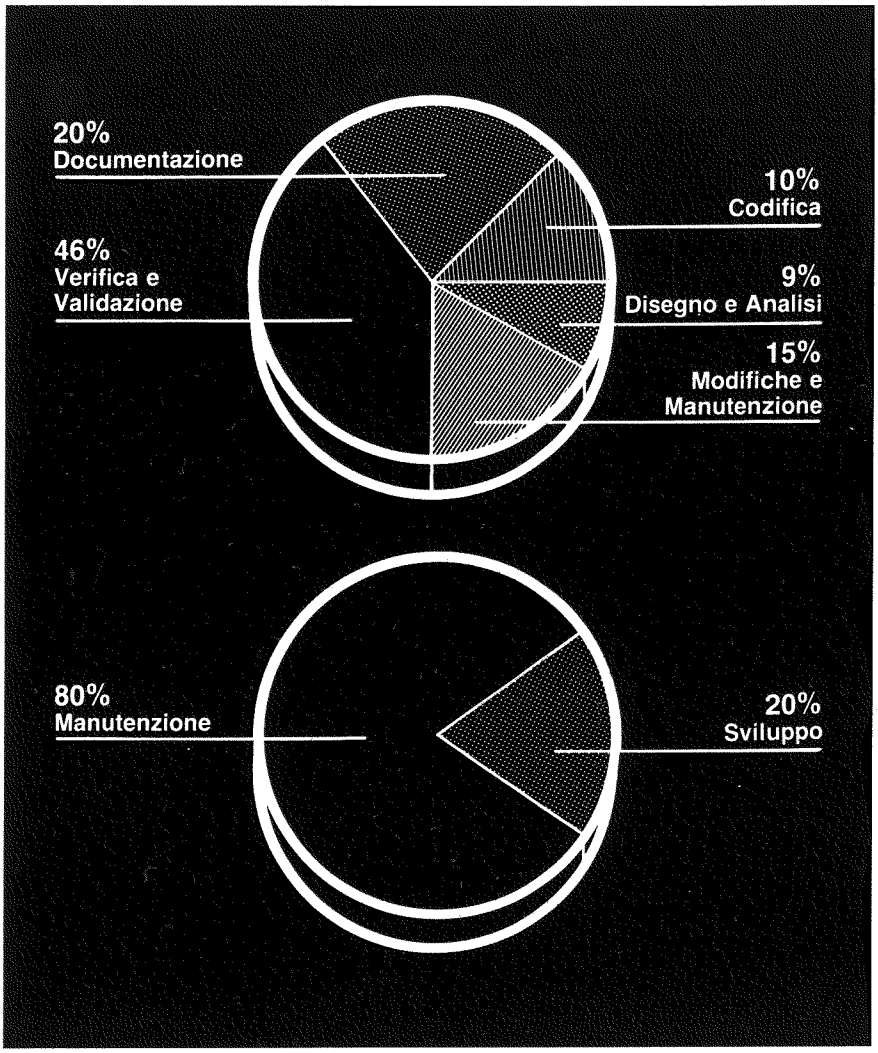
In realtà tale termine risulta fuorviante se associato al significato usuale, in quanto poco ha a che fare con i problemi di mantenere efficiente un prodotto.

La manutenzione del software viene infatti classificata in:

- *correttiva,* orientata alla correzione degli errori trovati (copre circa il 20% del totale dei costi di manutenzione);
- *perfettiva,* orientata al miglioramento delle caratteristiche prestazionali e di manutenibilità;
- *adattiva,* orientata alla crescita del prodotto per le cambiate esigenze dell'ambiente in cui è inserito il prodotto; tale tipo di manutezione, che corrisponde in tutto e per tutto a un'attività di sviluppo che parte da un prodotto già esistente anzichè partire da zero, copre la maggior parte dei costi di manutenzione.

Quindi gli obiettivi economici della produzione di software, a parità di qualità, devono essere orientati a ridurre i costi di manutenzione, piuttosto che non quelli di sviluppo. Nonostante che la manutenzione veda caratteristiche rilevantemente vicine a quelle dello sviluppo, numero-

Fig. 2 - Distribuzione dei costi nella produzione di software.



se sono le fonti di costo aggiuntivo che possono essere ridotte dal controllo di qualità:

- innanzitutto, la manutenzione correttiva è significativamente ridotta se il numero di errori presenti nel prodotto rilasciato è ridotto, e, pur non considerando primario in tal senso il ruolo del controllo di qualità, è indubbio che tale attività reperisce errori e quindi ne riduce la presenza al rilascio;
- inoltre il controllo di qualità interagisce significativamente con l'attività di sviluppo, inducendo un certo insieme di attributi nel prodotto sviluppato, come la leggibilità, la tracciabilità, e così via, che inevitabilmente rendono il prodotto più modificabile e manutenibile.

Quindi non si pone il problema di quando sia conveniente da un punto di vista del rapporto costo/beneficio prevedere un'attività di controllo di qualità: questa è sempre conveniente; si tratta piuttosto di rendere questa attività il meno costosa possibile di per sè, ad esempio attraverso metodi, strumenti e automazione, e di essere in grado di misurare la quantità di questa attività svolta, allo scopo di determinare il momento in cui essa debba considerarsi sufficiente, e quindi possa essere interrotta.

Le affermazioni fatte contrastano con uno scenario reale in cui l'attività di controllo di qualità per il software è effettuata raramente e, il più delle volte, malamente. Di fatto gli ambienti in cui la convinzione della inevitabilità di un tale lavoro è assoluta sono molto pochi, e spesso si ritiene che il costo aggiuntivo del testing sia inaccettabile per il mercato.

Spesso quindi, o addirittura quasi sempre, il controllo di qualità viene lasciato al progettista stesso dello sviluppo, che, con qualche prova non sistematica e raffazzonata, acquieta la coscienza del responsabile del progetto.

Proviamo invece ad esaminare cosa succede quando il controllo di qualità non viene effettuato in modo organico e sistematico.

a. Alcuni esempi

1. È ampiamente riportato e documentato il caso occorso pochi anni fa relativo a una sonda spaziale diretta verso Venere che, subito dopo la partenza da Cape Canaveral, ha mostrato una traiettoria scorretta che l'avrebbe portata a cadere su un'area densamente popolata, e che impose la distruzione in volo; analisi successive evidenziarono che il problema era legato a una inadeguata attività di controllo di qualità del

software, e in particolare un'istruzione FORTRAN del tipo

DO 10 I = 1,5

era stata scritta con un "." al posto della ","

DO 10 I = 1.5

facendo sì che il compilatore la interpretasse come un'istruzione corretta di assegnamento del valore reale 1.5 alla variabile DO10I dichiarata implicitamente.

2. Alcuni anni fa un centro di ricerca collegato all'European Space Agency partecipò a un progetto in cui un satellite rilevava, tra l'altro, la frequenza di raggi cosmici nelle zone esterne all'atmosfera terrestre; dopo una messa in orbita perfetta e una raccolta ricchissima di dati, i ricercatori dell'ente iniziarono a esaminare i valori inviati, ricavando interessantissimi risultati scientifici che pubblicarono puntualmente; dopo qualche mese un conteggio negativo da parte del satellite mise in allarme i tecnici: risultò che, per un errore software, tutti i dati giunti fino a quel momento erano scorretti e privi di alcuna utilità; non è noto se i risultati pubblicati siano stati successivamente denunciati o se la cosa sia stata messa a tacere.

3. È noto il diffuso fenomeno del "computer crime": si calcola che la rendita di uno di tali crimini sia da 10 a 50 volte quello di una rapina a mano armata, così come si calcola che solo il 15% di essi sia denunciato, in quanto la cattiva immagine che il fatto potrebbe arrecare al "derubato" (spesso una banca) causerebbe danni superiori ai benefici della denuncia e persecuzione; nonostante che spesso tali crimini vengano compiuti attraverso interventi dolosi nel software da parte di addetti o da uso doloso di funzioni permesse, sono numerosi i casi in cui l'intervento è fatto dall'esterno, sfruttando errori o mancati controlli che permettono di inserirsi in reti di calcolatori e manipolare o attingere a informazioni riservate. Famosi sono i casi di accesso e manipolazione dei cambi ufficiali delle valute nei calcolatori della Borsa di Londra ad opera di alcuni studenti dotati di un semplice home computer, e quello dell'accesso, in situazioni analoghe, al CRAY 1 della difesa militare francese del luglio '86.

4. Alcuni anni fa un satellite metereologico francese, per un errore software, invece di interrogare alcuni dei palloni sonda ad alta quota per ricevere informazioni, inviò a 72 dei 140 palloni l'ordine di autodistruzione.

5. Durante le missioni Apollo, numerosi errori software furono trovati: in una di esse una sequenza di comandi operatore non prevista distrusse informazioni registrate in memoria centrale, fortunatamente non critiche; in altre, varie volte gli astronauti furono costretti a distogliere la loro attenzione da attività previste nella missione, per intervenire manualmente sul comando.

6. Nel 1977, durante una prova di controllo di sistemi di comunicazione, controllo e comando militari statunitensi, il Worldwide Military Command and Control System presentò errori nel 62% dei messaggi comunicati, mentre il Readyness Command risultò non operante l'85% dei casi.

In tutti questi casi è evidente come il danno arrecato sia di portata inconfrontabile con i costi di un'attività di controllo di qualità, comunque costosa questa possa rivelarsi: nei primi due casi, infatti, il danno è stato pari al costo dell'intero progetto (hardware, software, attrezzature) inevitabilmente di gran lunga superiore a un costo che al più sarebbe stato pari al 50% del solo costo del software; nel terzo caso i rischi sono addirittura incalcolabili, in quanto possono essere riprodotti e moltiplicati nel tempo.

b. Una misura reale dei costi del controllo di qualità del software

L'attività di controllo di qualità consuma una fetta significativa dei costi dello sviluppo, in particolare per i progetti di dimensione medio grandi e collegati a rischi economici o umani; i seguenti dati sono riportati in letteratura:

Progetto	*% QC su sviluppo*
Sage	47%
Vaval Tactical Data System	50%
Gemini	47%
Saturn 1	44%
Apollo	80%

È tuttavia una misura diffusa il fatto che, in progetti di dimensioni medio-piccole o di non eccezionale criticità, sia accettabile uno sforzo di testing che arriva a coprire circa il 30% dei costi.

Tale sforzo si dimostra influire sulla qualità del prodotto, riducendo i costi di manutenzione correttiva e aumentando la manutenibilità (ovvero riducendo i costi di manutenzione perfettiva e adattiva).

Facendo l'ipotesi di un progetto che costa 30 di sviluppo e 70 di manutenzione, possiamo fare i seguenti conti, significativi anche se grossolani, sul risparmio indotto dalla eliminazione dell'attività di controllo di qualità:

- riduzione dei costi di sviluppo (-30% dello sviluppo, ovvero del 30% del costo totale di progetto): -10
- riduzione dei costi di manutenzione adattiva per assenza di qualificazione (-30% del 60% della manutenzione adattiva e perfettiva, che può essere stimata intorno al 70% del progetto: quindi il 30% del 60% del 70%): -13
- aumento dei costi di manutenzione correttiva (raddoppio della quota cioè due volte il 20% del 70%): +28
- aumento della manutenzione perfettiva (aumento del 50% della quota): +5

È evidente che, da tali calcoli, risulta che non fare controllo di qualità costa più che farlo, ma anche se si volessero rifiutare i numeri indicati modificandoli verso una tesi opposta, risulta evidente che il risparmio ottenuto sarebbe dell'ordine al più di qualche % dei costi totali, a fronte di rischi molto elevati.

c. I costi di correzione degli errori

Di fatto i valori forniti nella rozza stima precedente sono ottimistici,

nel senso che non fare il controllo di qualità ha costi molto superiori a quelli indicati; la fig. 3 illustra l'andamento dei costi di correzione di un errore a fronte del momento in cui esso viene scoperto [2].

Da essa si vede come, fatto 1 il costo di correzione di un errore individuato durante la definizione dei requisiti, la correzione dello stesso errore scoperto durante l'uso, quando cioè il prodotto è ormai rilasciato, può costare da 200 a 500. Ciò può avere la seguente spiegazione intuitiva:

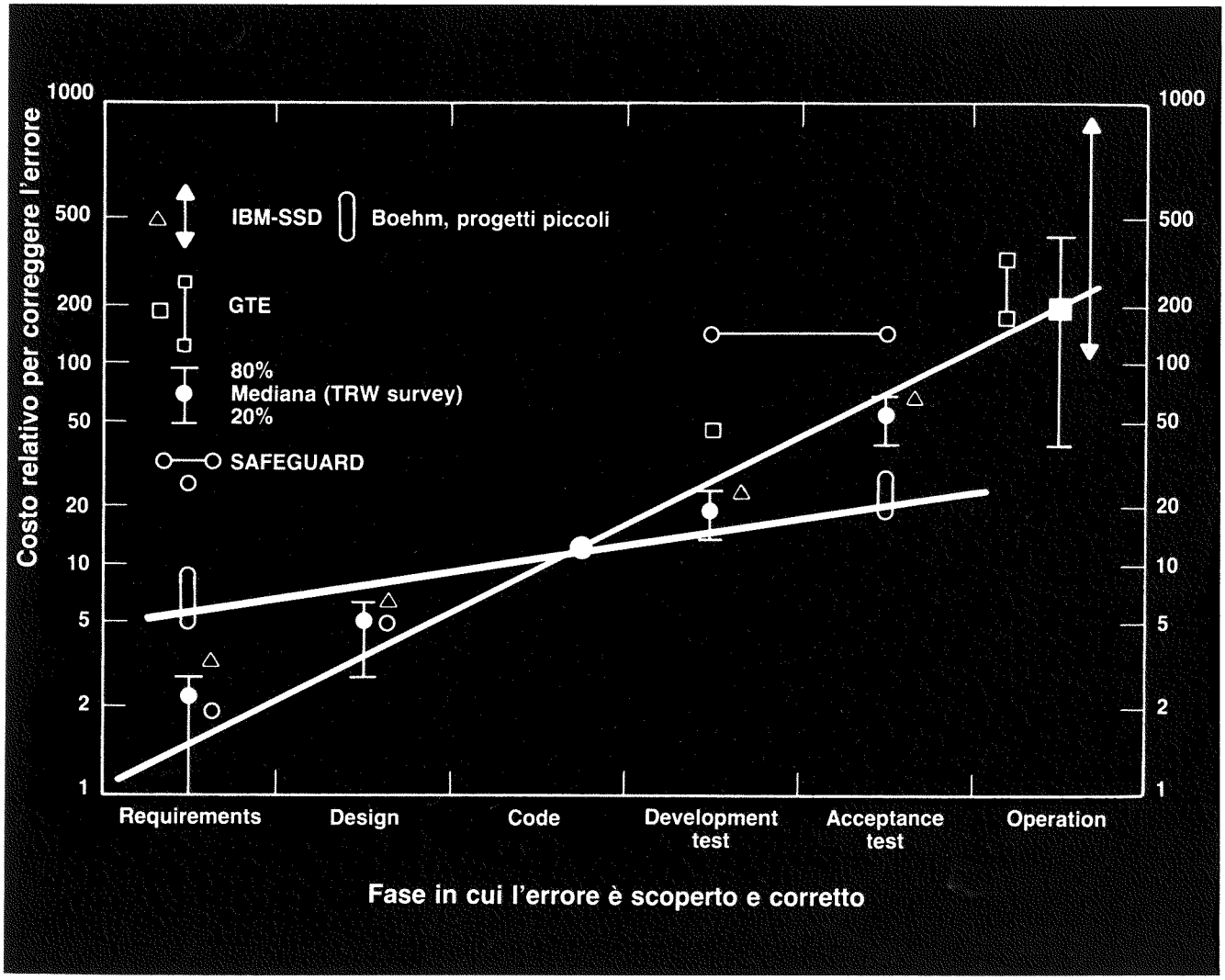
- un errore scoperto durante la fase di codifica è, tipicamente, un errore molto localizzato: un'interfaccia non congruente col progetto, un calcolo non coerente con la specifica, ecc.; la sua correzione è legata alla modifica locale;
- un errore scoperto durante la fase di qualificazione è tipicamente legato al fatto che il prodotto, così come è stato costruito, non rispetta le specifiche funzionali (tale è infatti il punto di riferimento della qualificazione); ciò può essere legato a una errata architettura, e comunque l'intervento non è in genere locale, ma coinvolge aspetti più profondi del sistema;
- un errore scoperto durante l'uso rischia di essere la conseguenza di un'errata analisi, che ha specificato e progettato un sistema non adeguato alle necessità; è chiaro che un intervento in tal caso è inevitabilmente più gravoso.

A ciò va anche aggiunto il fatto che, per quanto la documentazione del prodotto sia completa e adeguata, è sempre necessario un significativo sforzo di studio delle caratteristiche fisiche dei programmi e delle motivazioni che hanno portato a determinare scelte realizzative che cresce quanto più tempo passa dal momento in cui il prodotto è stato realizzato.

Le caratteristiche del diagramma di figura 3 evidenziano che il controllo di qualità non deve essere un'attività che interagisce con il progetto solo alla fine del completamento del prodotto, ma che deve accompagnare con analisi critiche e revisioni ogni singola fase, valutando l'adeguatezza dei prodotti di ciascuna delle fasi di sviluppo.

Le considerazioni fatte sono ulteriormente esasperate dalla fig. 4 che illustra una reale esperienza di mi-

Fig. 3 - Andamento dei costi di correzione di un errore in funzione della fase in cui esso viene reperito [1].



sura di costi e di risparmi di costi attraverso un controllo di qualità sistematico e portato avanti con l'uso di strumenti automatici: l'esperienza, condotta presso i laboratori della Hughes Aircraft Company, evidenzia che l'anticipazione della scoperta di errori permessa dall'uso di una qualificazione sistematica ha ridotto a meno di un sesto i costi di correzione degli stessi [4].

d. Perché il controllo di qualità non si fa

A questo punto ci si chiede come mai i produttori di software siano così poco accorti, in media, da non adottare sistematicamente tecniche di testing; diversi sono i motivi riscontrabili; tra questi, i seguenti:

* i costi di un controllo di qualità sistematico sono facilmente evidenziabili, mentre i costi derivati dalla sua assenza non sono facilmente misurabili, e comunque non sono in genere soggetti a una rilevazione (per inciso, raramente negli ambienti di produzione di software si effettua una precisa rilevazione dei costi di produzione, per i quali esistono, nella migliore delle ipotesi, semplicemente delle stime preventive, in genere largamente al di sotto di quelle consuntive); così la scelta è per evitare un costo evidente immediato che uno non noto futuro;
* la scarsa sensibilità e preparazione dell'utente finale, che si muove tra prodotti che, sulla carta sembrano fornire lo stesso servizio, ma con costi drasticamente diversi, e l'assenza di una qualsiasi garanzia sui prodotti fa sì che gli aspetti di concorrenza, in particolare sui mercati di livello più basso o di prodotti con scarsa tradizione, prendano il sopravvento; spesso una piccola software house è costretta a tagliare i costi di qualificazione perché la sua "contabilità" risulterebbe avere un prezzo fuori mercato.

Di fatto una situazione simile è ancora possibile per una certa classe di prodotti che non esigono alta affidabilità e correttezza (ad esempio una procedura di contabilità), e che, in particolare, possono contare sul fatto che la qualificazione venga "sopportata" dal cliente; questi, reperendo e segnalando errori, via via fa evolvere il prodotto a un adeguato grado di stabilità che ne permette una distribuzione più ampia.

In effetti, se è assolutamente inevitabile affrontare costi altissimi (si pensi all'80% del progetto Apollo) per qualificare prodotti che presentano alti rischi in caso di malfunzionamento, non è sensato un analogo investimento per prodotti commer- ciali tipici per i quali, come si è già detto, un investimento tipico raggiunge il 30% dei costi di sviluppo.

Fig. 4 - Vantaggi economici di un'attività di qualificazione sistematica di prodotti software.

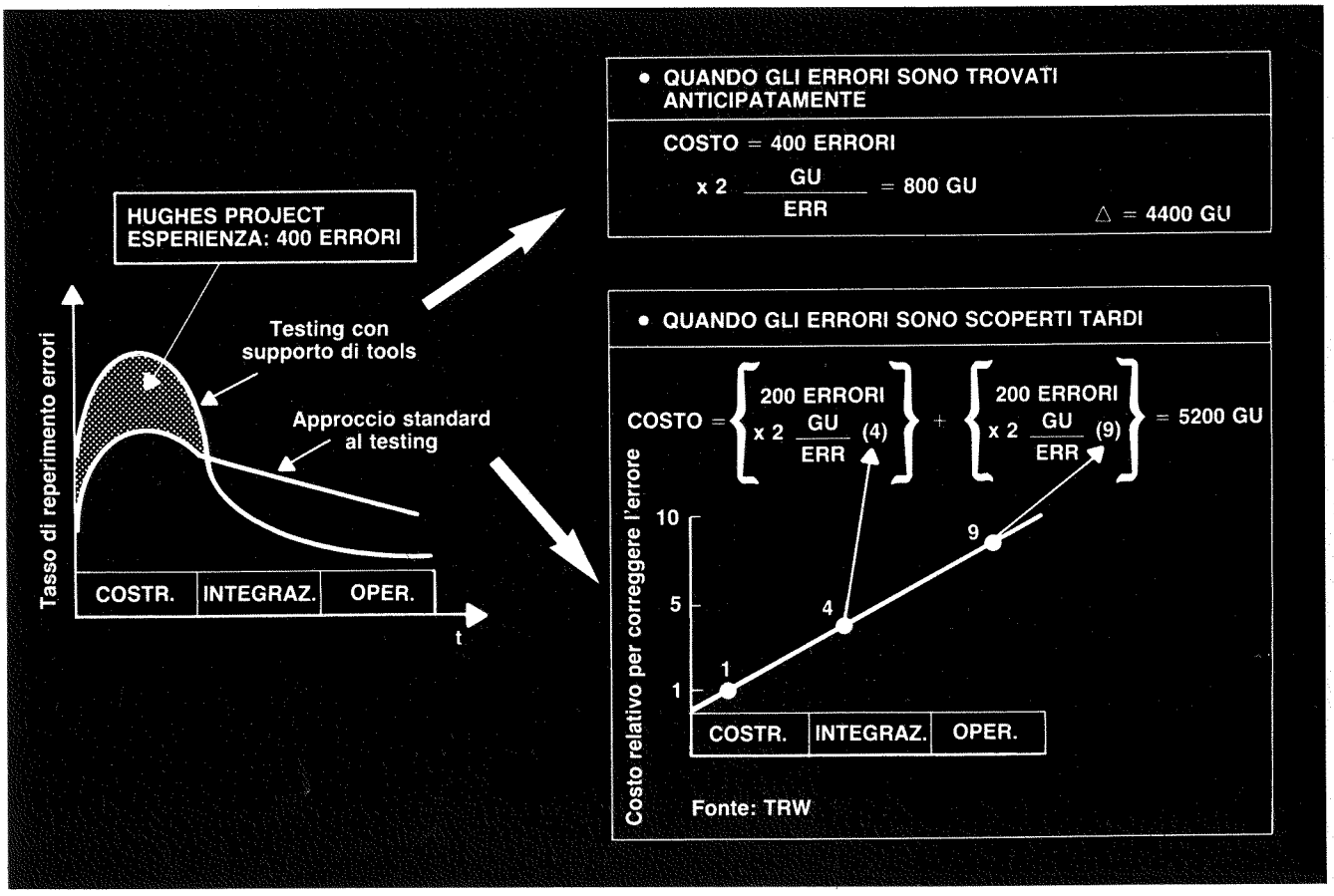
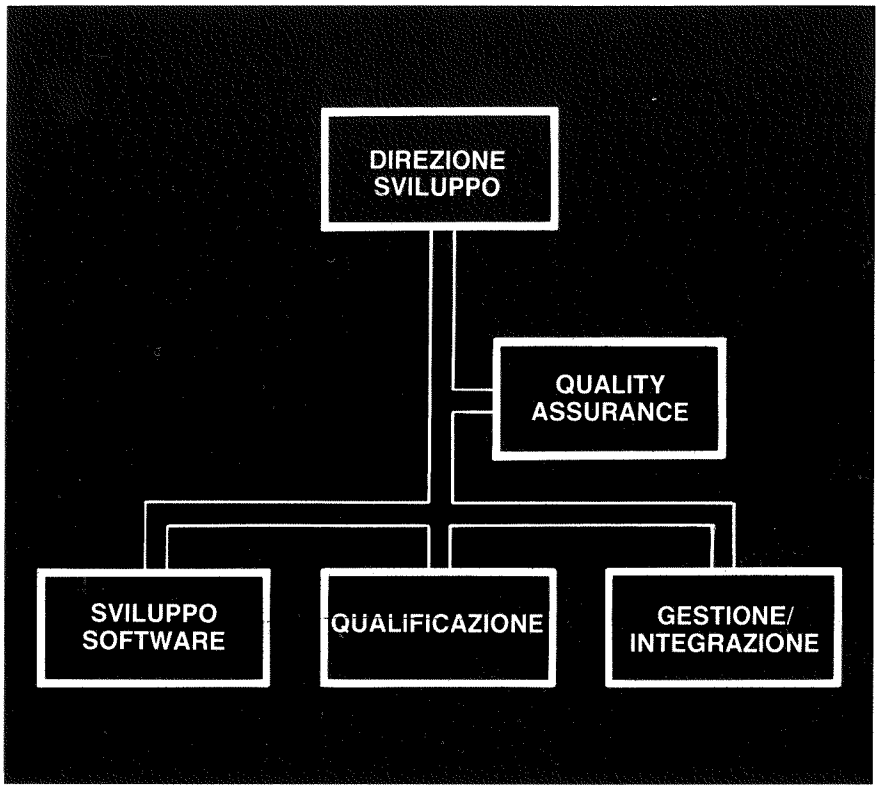


Fig. 5 - Organigramma di un'organizzazione aziendale che sviluppa software.



3. Organizzazione del lavoro e controllo di qualità del software

Poichè l'attività di controllo di qualità deve essere collocata nell'ambito dello sviluppo in modo che si integri correttamente con questo, è opportuno esaminare quali strutture organizzative e funzionali siano adeguate a risolvere questo tipo di problema.

La figura 5 illustra un'organigramma molto schematico che colloca l'attività di controllo di qualità: in esso si vedono tre entità che rispondono gerarchicamente al responsabile dello sviluppo e una che è invece di staff a quest'ultimo. Esaminiamoli uno per uno.

- Lo **sviluppo di software,** che, a partire dalle specifiche, produce i vari moduli e documenti intermedi, che dovranno poi essere archiviati, gestiti, e che saranno il punto di partenza per l'integrazione del prodotto finale.
- La **gestione/integrazione,** che prende in carico i moduli e i documenti sviluppati, e provvede a catalogarli, gestirli, integrarli, evitando modifiche non autorizzate, e ineragisce con il controllo di qualità per la qualificazione.
- Il **controllo di qualità,** che effettua una qualificazione dei vari moduli e del prodotto globale, compreso le eventuali revisioni dei documenti, interagendo con la gestione/integrazione per avere il materiale da qualificare e con il gruppo di sviluppo per la notifica di eventuali anomalie.
- **L'assicurazione di qualità,** invece, è l'ente preposto a garantire tutte le premesse per una qualità a *priori* del prodotto, predisponendo le strutture organizzative, le prassi operative, i metodi e gli strumenti; spesso l'accezione di questo termine è differente, e si intende con "assicurazione di qualità" l'attività volta a verificare che tutte le attività di qualificazione, verifica e validazione su un progetto siano state adeguatamente svolte; noi preferiamo vedere tale attività come un compito specifico del gruppo di controllo di qualità.

Esaminando la struttura organizzativa da un punto di vista funzionale piuttosto che organigrammatico, possiamo osservare la fig. 6.

Da essa si vede che:

- il gruppo di **sviluppo** è "proprietario" di un *archivio di sviluppo,* in cui sono presenti tutti i documenti e moduli di lavorazione, sotto il controllo diretto dei progettisti, e il cui stato non è assolutamente stabile nè garantito; il gruppo interfaccia con una funzione di **pianificazione,** che è una funzione di direzione, per definire, concordare e aggiornare piani di sviluppo e rapporti di avanzamento; interfaccia invece con il gruppo di Gestione/Integrazione per la consegna di ciò che può essere considerato finito (documenti o programmi);
- il gruppo di **integrazione/gestione** è proprietario di un *archivio "congelato",* che contiene tutti i documenti e i moduli che non possono essere modificati se non a fronte di rigorose procedure di "change control", così da garantire la stabilità di tutti i prodotti verso la qualificazione, verso l'attività di integrazione e di rilascio; il gruppo fa riferimento a piani di attività che coincidono con quelli dello sviluppo, e cede al controllo di qualità l'accesso a moduli e a documenti per la loro qualificazione e verifica;
- il gruppo di **controllo di qualità,** facendo riferimento ai soliti piani,

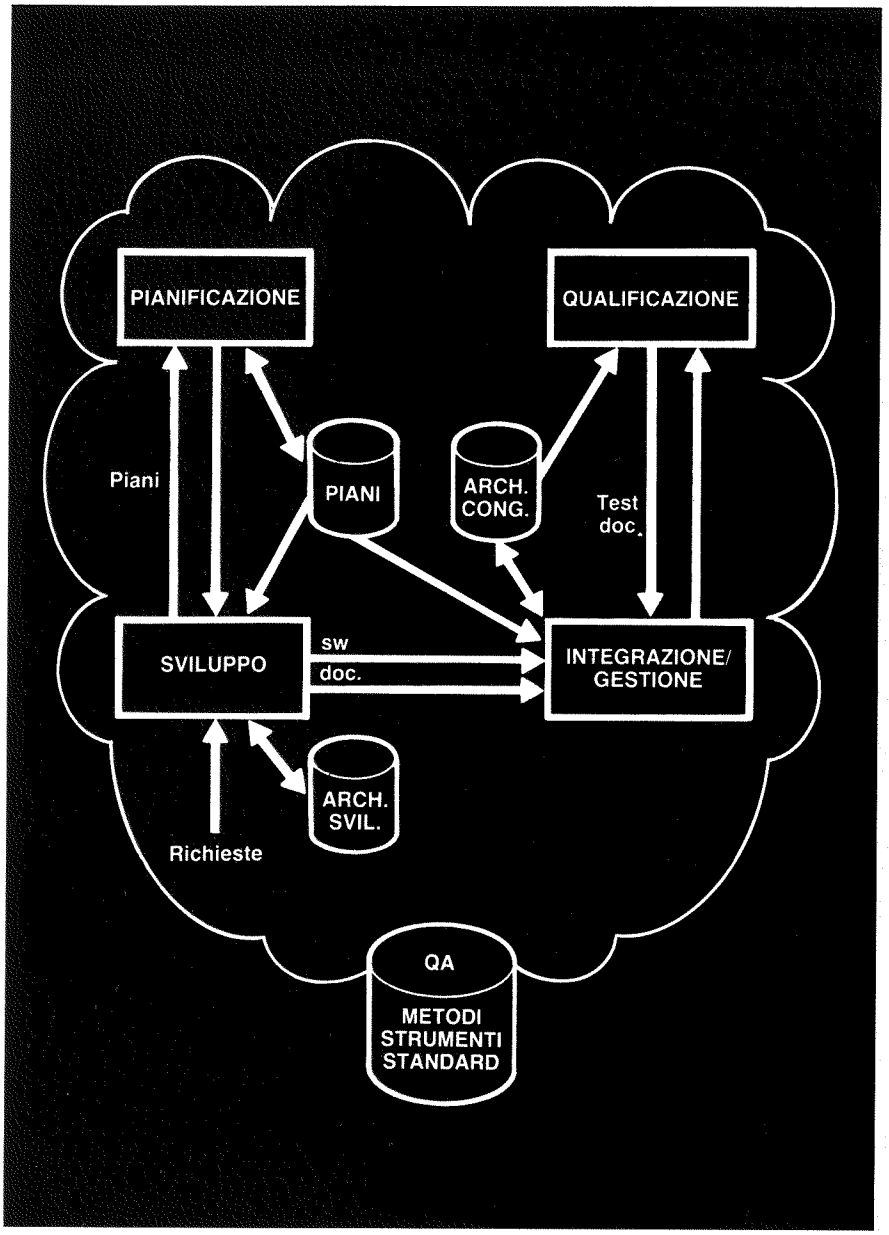


Fig. 6 - Organizzazione funzionale di una componente aziendale di sviluppo di software.

produce test per la qualificazione e rapporti di qualità e di revisione di prodotti e documenti, che vengono gestiti sempre dal gruppo di Gestione/Integrazione; tali rapporti sono alla base del giudizio di rilasciabilità del prodotto finale;

- tutti i gruppi indicati sono "immersi" in un ambiente che si muove con regole, procedure, prassi operative e strumenti definiti o forniti dall'**assicurazione di qualità.**

Una tale visione funzionale è troppo semplicistica dal punto di vista di una grande organizzazione, ed è troppo complessa per una piccola organizzazione; tuttavia lo scopo della trattazione non è quello di suggerire una reale organizzazione del lavoro quanto quello di evidenziare un certo insieme di principi; una piccola organizzazione potrà, su un piccolo progetto, avere il 30% di una persona che rappresenta il controllo di qualità, costituendo per il restante 70% il progettista di sviluppo, che, ben lungi dal sottoporsi a procedure burocratiche di "change control", avrà cura di avere sempre due sole versioni di un modulo (una in lavorazione e una all'ultimo stato stabile), e che saprà impersonare differenti ruoli con diversi atteggiamenti passando da progettista a controllore.

4. Il problema della certificazione

Finora abbiamo parlato di "qualificazione del "software", come di quell'attività interna a un ambiente di produzione (o comunque dipendente da questo dal punto di vista di responsabilità e obiettivi), il cui compito è quello di accertare la congruenza del prodotto finale con le aspettative iniziali.

Quale tipo di garanzia per l'eventuale acquirente è insito in una tale attività? In realtà la conoscenza della presenza di buone organizzazioni di qualificazione interne all'ambiente di produzione è di per sè un elemento che accresce la confidenza nel prodotto da parte dell'acquirente,

ma, da un punto di vista strettamente legale, la garanzia è pressochè nulla, tanto che clausole di rimozioni di errori trovati dopo il rilascio sono esplicitate di volta in volta in ogni contratto, e, in genere, vengono previsti contratti di manutenzione per la rimozione di errori successivi a un certo periodo dopo la installazione.

Di fatto, i prodotti software di larga diffusione specificano nelle clausole contrattuali che i produttori non sono responsabili di qualunque evento causato dall'uso corretto o scorretto del loro prodotto, ancorchè tali eventi possano derivare da errori presenti nel prodotto stesso.

Nasce quindi il problema della *certificazione*, ovvero dell'autenticazione delle attività di qualificazione svolte dal produttore stesso o svolte da altri enti autorizzati.

Uno scenario organizzativo per la certificazione di un prodotto presuppone:

- una *normazione di prodotto* che precisi i requisiti di una specifica classe di prodotti, eventualmente insieme a requisiti sulle relative modalità di lavorazione
- un *produttore* in grado di mostrare la sua organizzazione produttiva e di documentare le prove effettuate sul prodotto
- un *laboratorio di prova* in grado di pronunciarsi sulla valutazione di un prodotto, sia basandosi su ispezioni del processo produttivo, sia avvalendosi di documenti di prova del produttore, sia con prove in proprio
- una *normativa di accreditazione* di un laboratorio di prova, affinchè questo venga dichiarato competente a pronunciarsi sulla valutazione di un determinato prodotto
- un *ente di certificazione* che, sulla base del pronunciamento di un laboratorio di prova accreditato, certifichi la *conformità ai requisiti* del prodotto sviluppato.

Come si vede la situazione è tuttaltro che semplice, e le difficoltà di poter definire chiari requisiti per un qualunque prodotto software mostrano quanto sia ancora lontana un'attività di certificazione dello stesso: è tuttavia pensabile anche a breve una sorta di "attestato" che dichiari che un certo prodotto è stato sviluppato "a regola d'arte", indicando così che, stanti i limiti ancora impliciti nella tecnologia di produzione del software, ogni tecnica, metodo o strumento adatti all'ottenimento di un prodotto di adeguati requisiti sono realmente stati impiegati, e che quindi qualunque errore presente nel prodotto è da considerarsi dovuto a "caso fortuito".

Di fatto, l'atteggiamento presente in forniture software per enti governativi e militari americani, tra i più avanzati in quest'ambito, va in tale direzione: i produttori vengono accreditati allo sviluppo di software per tali enti sulla base della loro organizzazione della produzione, e l'accettazione di un prodotto è soggetta a un "piano di qualità" che esamina con dettaglio le modalità di sviluppo e di qualificazione relative alla singola fornitura.

Riferimenti bibliografici

[1] B.W. BOEHM, *Developing Small Scale Application Software Products: Some Experimental Results,* Proc. IFIP 8th, Oct. 1980.

[2] B.W. BOEHM, *Software Engineering Economics,* Prentice-Hall, 1982.

[3] A. CICU, M. MAIOCCHI, *Il testing dei prodotti software: il punto di vista del responsabile di produzione,* Convegno Aica 1982, Padova.

[4] M.S. DEUTSCH, *Software Verification and Validation: Realistic Project Approach,* Prentice-Hall, 1982.

[5] G. W. MYERS, *The Art of Software Testing,* Wiley, 1979.

Microelettronica per la sicurezza: le *smart card*

ANTONIO BRIOSCHI,
ALESSANDRO GROSSI

Business Unit Card Systems
Bull Italia
Milano

1. Introduzione

Il rapporto esistente tra uomo e tecnologia è tale che ad ogni innovazione corrispondono delle reazioni comportamentali e di sistema di eguale portata.

La ricerca infatti dà come risultato dei prodotti che, in buona parte dei casi, vengono industrializzati ed entrano a far parte del sistema modificandone la struttura stessa.

L'elettronica applicata alla vita quotidiana, in particolare, ha portato ad una vera e propria rivoluzione nei costumi e nelle abitudini dei cittadini.

Oggi, alla soglia degli anni novanta, stiamo vivendo un periodo in cui le innovazioni del settore sono rapidissime, e, di conseguenza, diviene quasi impossibile riuscire a stare al passo coi tempi.

Restringendo il campo di esame al tema della sicurezza logica e fisica, conviene fare una veloce carellata sulle tecnologie passate ed attuali, prima di scendere all'esame delle caratteristiche delle soluzioni basate sulle *smart card.*

2. Breve storia delle "carte"

Le prime carte personali, nate per permettere l'identificazione del portatore, erano costituite proprio di materiale cartaceo.

Negli anni cinquanta e sessanta, poi, si è iniziato ad utilizzare le tessere plastiche che portavano incisi, o in rilievo, gli elementi identificativi del titolare.

Con la diffusione degli strumenti elettronici e dei punti di accesso remoti ai sistemi per l'elaborazione dei dati, si è sentita la necessità di permettere un riconoscimento automatico del portatore di carte plastiche.

Ecco così nascere le carte con la *striscia magnetica* e con i *codici ottici.* Queste soluzioni hanno avuto una naturale evoluzione, che ha visto prevalere la tecnologia della striscia magnetica per moltissimi motivi.

Queste carte vengono dette *"passive"* in quanto possono soltanto archiviare dei dati ed hanno una capacità di memorizzazione molto limitata.

Con la miniaturizzazione della componentistica elettronica si è arrivati, negli anni ottanta, alla possibilità di produrre dei *"chip",* o meglio dei *"microchip",* di dimensioni così contenute da poter essere inseriti nello spessore di una tessera plastica rispondente alle norme ISO che regolano questa materia.

Parallelamente, poi, si è sviluppata la tecnologia *laser* che ha consentito di produrre dei sottilissimi film codificabili, più o meno come avviene per i supporti dei compact disc e dei dischi ottici.

Attualmente, per i non esperti, le differenze fra le varie tecnologie non sono apprezzabili con molta facilità.

Le carte laser, prima di tutto, sono di tipo passivo e servono per memorizzare una grande massa di dati; per quanto riguarda la sicurezza delle informazioni contenute e degli accessi, però, ci si deve affidare essenzialmente ad apparati esterni alla carta.

Le carte a chip si dividono in due grandi categorie: quelle *a memoria,* dette anche a *"microchip",* e quelle dotate di un *microprocessore,* dette anche *"smart card".*

Le carte a memoria, anche se del tipo con logica cablata, sono prodotti di basso costo, utilizzati in genere per applicazioni nelle quali non è necessaria una elevata sicurezza interna.

La componentistica di base può essere acquistata con facilità ed i produttori di microchip diffondono pubblicamente le informazioni su come sono costruiti e su come si programmano i loro chip.

La sicurezza delle carte di questo tipo, quando è richiesta, si basa su apparati e dispositivi esterni alla carta.

Una applicazione tipica di carte a microchip è quella per le carte a decalaggio dei telefoni pubblici francesi.

Queste carte in genere vengono gettate dopo l'uso.

Ben diverse sono le carte a microprocessore o smart card, che includono, al loro interno, un chip dotato non solo di memoria e registri, ma soprattutto di un vero microprocessore programmabile, che non si limita ad alzare una barriera tra l'interno della carta ed il mondo esterno, ma la rende un *elemento attivo* del sistema.

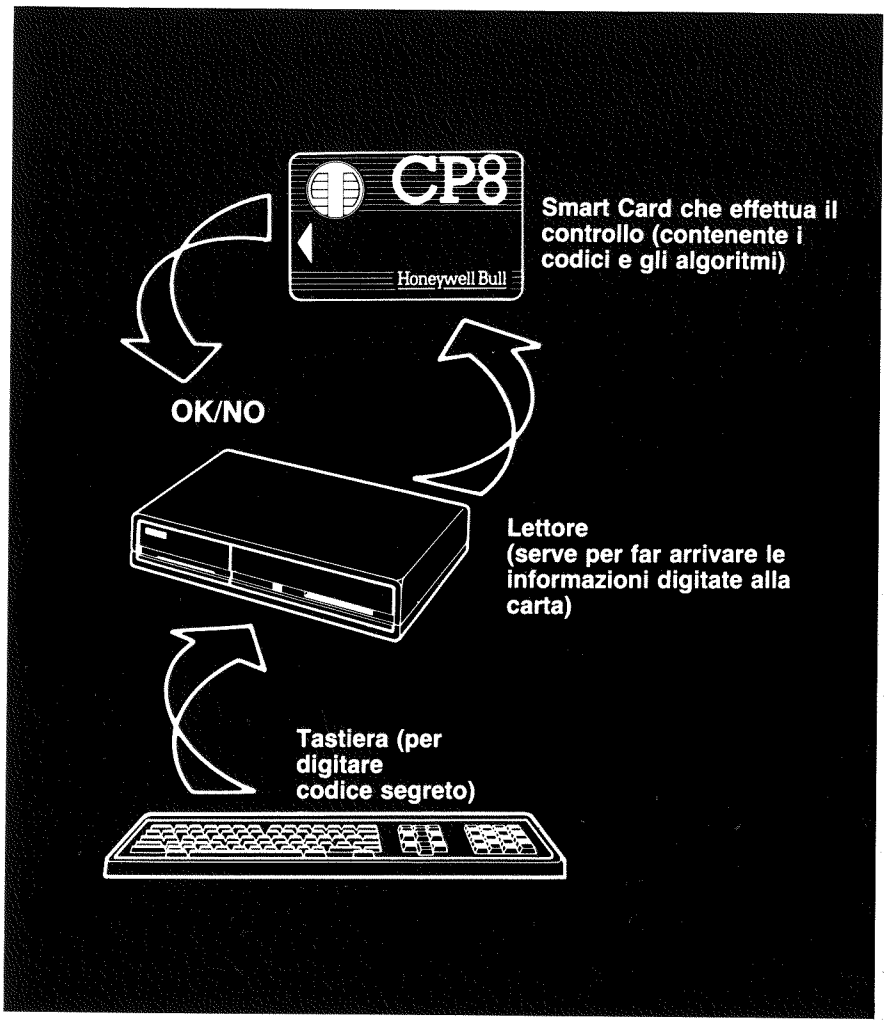


Fig. 1 - Identificazione dell'utente.

L'essere un vero e proprio computer, infatti, è una caratteristica di base delle *smart card,* che permette di raggiungere livelli di sicurezza non ottenibili con altri tipi di carte personali.

Un altro elemento di valutazione della soluzione tecnologica è dato dal fatto che la *smart card* CP8, la prima ed ancora oggi l'unica carta a microprocessore prodotta su scala industriale, non soltanto ha dimostrato la sua validità in moltissimi esperimenti, ma è ormai definibile come un prodotto maturo, utilizzato in sistemi operanti in effettivo e su grande scala.

Disporre quindi di un *elemento attivo ed inviolabile,* che può essere inserito nel portafoglio, cosa significa ai fini della definizione di un sistema di pagamento elettronico o per la protezione delle informazioni?

Esaminiamolo analizzando la problematica della sicurezza logica.

3. Le *smart card*

Si può affermare che, disponendo di strumenti come le *smart card* CP8, il compito degli informatici per proteggere le reti, i sistemi e le banche dati da accessi fraudolenti, diviene oggi più facile e meno oneroso; e questo senza venir meno ai requisiti di sicurezza, ma aggiungendo piuttosto un'ulteriore barriera ai tentativi di violazione, poichè le *smart card* CP8 sono un supporto leggero, discreto, induplicabile ed inaccessibile che ogni "responsabile" può avere con sè in qualsiasi momento ed in qualsiasi luogo.

Per quanto riguarda la sicurezza intrinseca durante la fase di assegnazione delle carte si può dire, con un esempio, che il funzionario di banca responsabile della personalizzazione delle carte utente dispone di una *smart card* CP8 con la quale attivare il processo, e di cui soltanto lui conosce la chiave di abilitazione, ma non il contenuto che sarà ricopiato sulle tessere da personalizzare.

Una volta resa disponibile una "referenza" induplicabile ed inaccessibile, diviene possibile sviluppare un sistema *sicuro* di accesso alle informazioni.

La *smart card,* quindi, costituisce il primo anello della catena, e la fase di identificazione assume la struttura illustrata in figura 1.

Lo sviluppo degli algoritmi di controllo della chiave utente o *"pin code"* è a cura del microprocessore inserito nella carta, che rivela all'esterno soltanto una segnalazione di "OK"; diviene così impossibile duplicare o riconoscere il "meccanismo logico" che sta alla base dell'identificazione.

Nel campo della sicurezza logica e della protezione dei dati, però, l'esigenza di assicurare la corretta trasmissione dei messaggi assume oggi un peso sempre maggiore.

Non è più sufficiente un semplice riconoscimento di chi opera al terminale, ma occorre piuttosto "autenticare", "certificare", generare le chiavi per crittografare i messaggi.

Per autenticare in modo sicuro chi opera ad un posto di lavoro informatico, si deve utilizzare un sistema inviolabile di domanda e risposta.

Per farlo, infatti, non serve mascherare il contenuto del messaggio, ma bisogna istituire un vero e proprio dialogo tra due elementi sicuri.

In questo caso il processo si svolge attraverso una carta "madre", presente presso la banca dati, ed una serie di carte "figlie" distribuite agli aventi diritto alle informazioni.

L'autenticazione, nel caso delle *smart card,* richiede l'esecuzione di una identificazione preventiva dell'operatore, che viene effettuata localmente con le modalità descritte in precedenza.

Ad esito positivo, e soltanto a questa condizione, il sistema si attiva ed il centro invia un numero casuale che può essere basato, ad esempio, sul momento stesso dell'invio (giorno, ora, minuto), ed è generato dalla carta madre.

La *smart card* CP8 "figlia" calcola ed emette un messaggio di risposta basato sui dati della carta e sul numero casuale ricevuto in precedenza; a partire da questo risultato, la carta madre è in grado di ricostruire i dati segreti utilizzati dalla carta figlia e stabilire se si tratta di una carta autorizzata o meno, ed autenticare quindi con sicurezza l'operatore.

Le *smart card* CP8, inoltre, garantiscono l'autenticazione "costante" durante tutte le transazioni, in quanto ogni tentativo di estrazione o sostituzione della carta "figlia" viene rilevato dal lettore locale e segnalato, provocando il blocco della stazione di lavoro.

Per ottenere una sicurezza ancora maggiore, infine, è sufficiente richiedere periodicamente una ridigitazione del codice segreto dell'operatore, provocando automaticamente una nuova autenticazione.

Il modello del trattamento è illustrato nella figura 2.

I messaggi, una volta identificato ed autenticato il loro emittente, posso- no richiedere una semplice verifica del contenuto, oppure, per maggior protezione, la crittografia del testo.

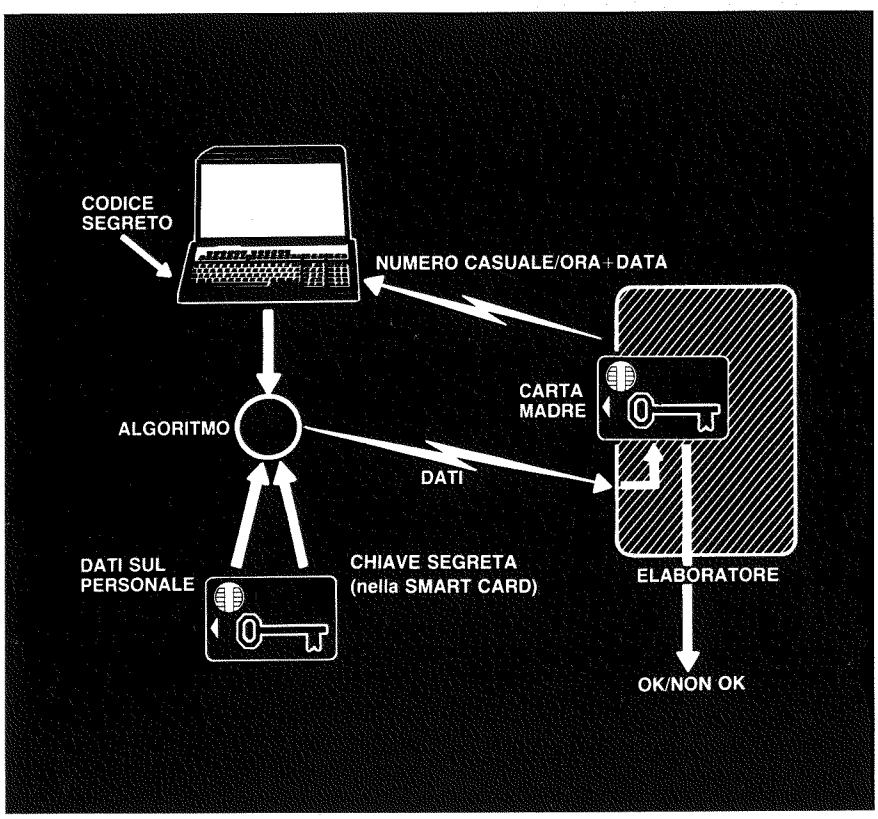
Nel primo caso è sufficiente predisporre un certificato emesso assieme al messaggio vero e proprio, che lo completa prima dell'invio in linea.

Una *smart card* può creare il certificato, utilizzando come variabile il testo del messaggio e come elemento di base il contenuto della carta stessa.

L'algoritmo di calcolo è conservato nelle memorie protette della *smart card,* e il risultato dell'operazione è ricalcolabile da parte dell'altra *smart card* presso la stazione ricevente.

La tecnica di certificazione può venire estesa, con gli opportuni adattamenti, anche al trasferimento di informazioni su supporti magnetici.

Fig. 2 - Processo di autenticazione mediante carta "madre".



La struttura di base della certificazione è esposta nella figura 3.

L'ambiente di utilizzo nel caso di una trasmissione di dati via linea è illustrato nella figura 4.

L'ambiente di utilizzo nel caso di una trasmissione di supporto magnetico (nastro, disco, diskette), è invece illustrato nella figura 5.

Nel caso si renda necessario proteggere il contenuto del messaggio anche dal semplice ascolto, l'unico sistema valido rimane quello della cifratura o crittografia, e in questo caso, ovviamente, la sicurezza del messaggio risiede essenzialmente nelle chiavi utilizzate per crittografarlo. Le tecniche utilizzate correntemente prevedono un frequente cambio delle chiavi stesse, al fine di prevenirne la fuga.

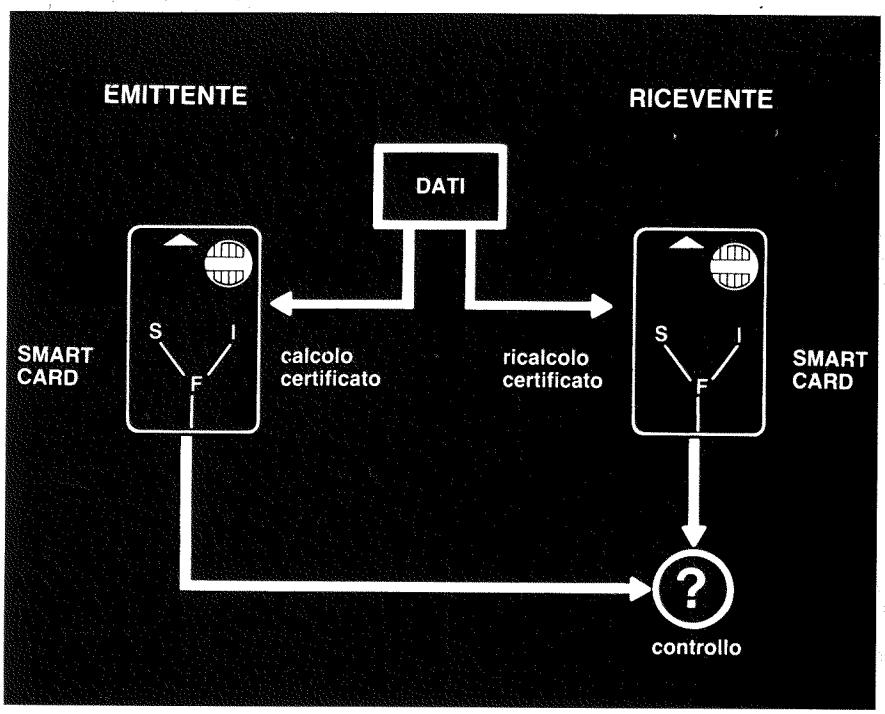


Fig. 3 - Struttura di base della certificazione.

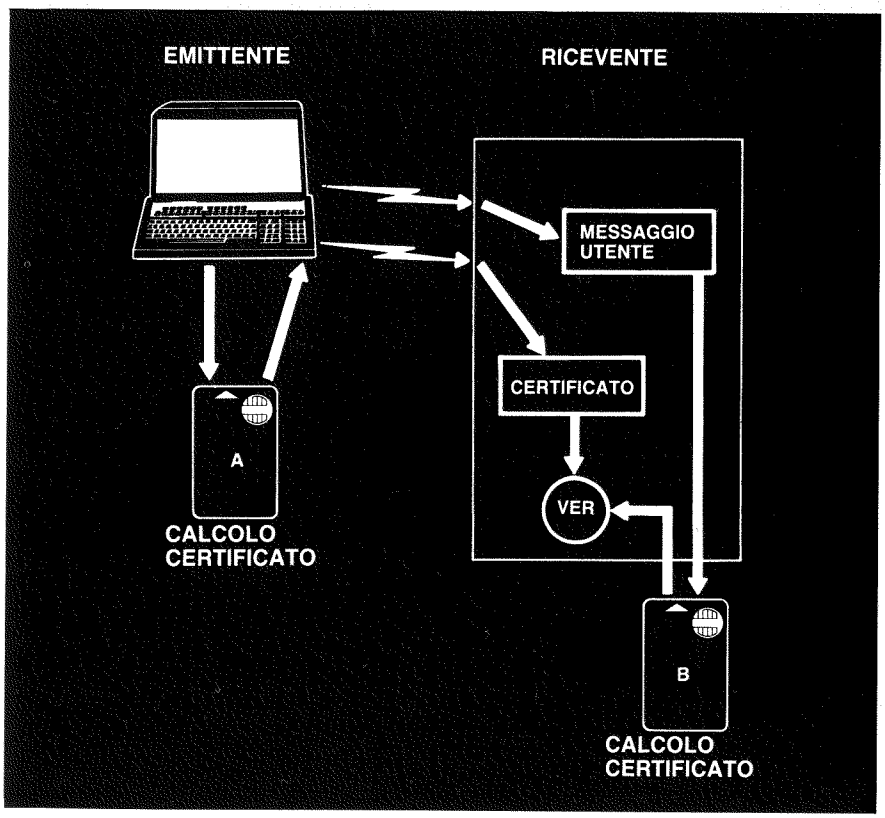


Fig. 4 - Ambiente di utilizzo nel caso di trasmissione dei dati via linea telefonica. (VER = verifica che certificato A uguale certificato B). Il messaggio utente viene introdotto dopo identificazione e autenticazione dell'utente.

Fig. 5 - Ambiente di utilizzo nel caso di trasmissione di supporto magnetico (L'asterisco (*) indica che la chiave banale può in questo caso essere trasmessa in chiaro).

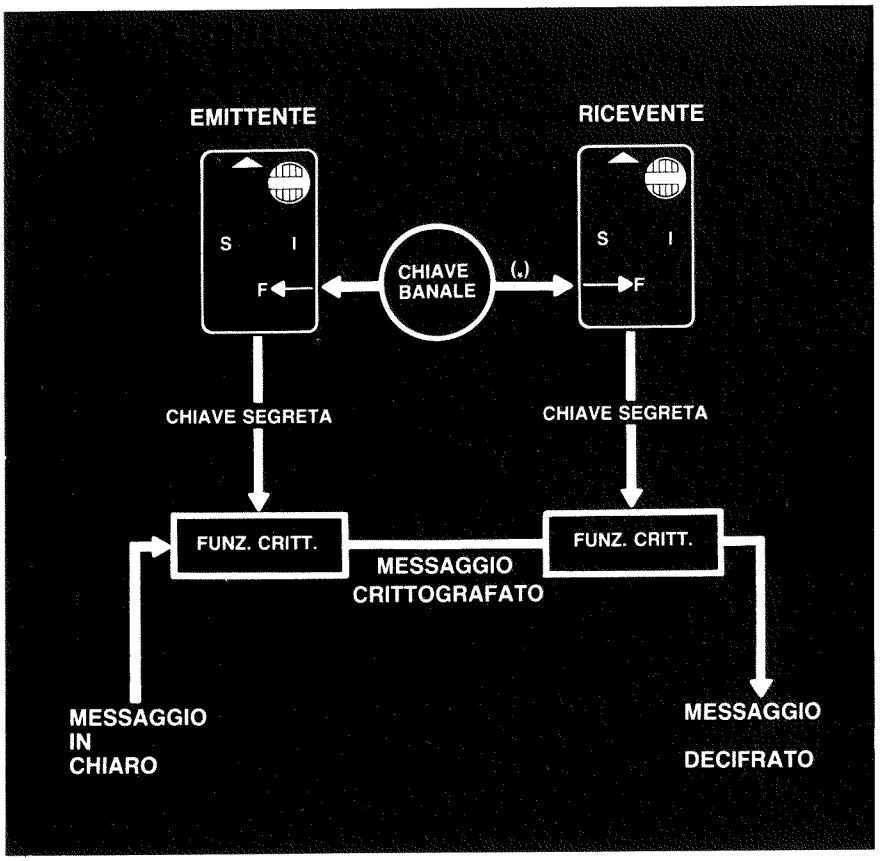
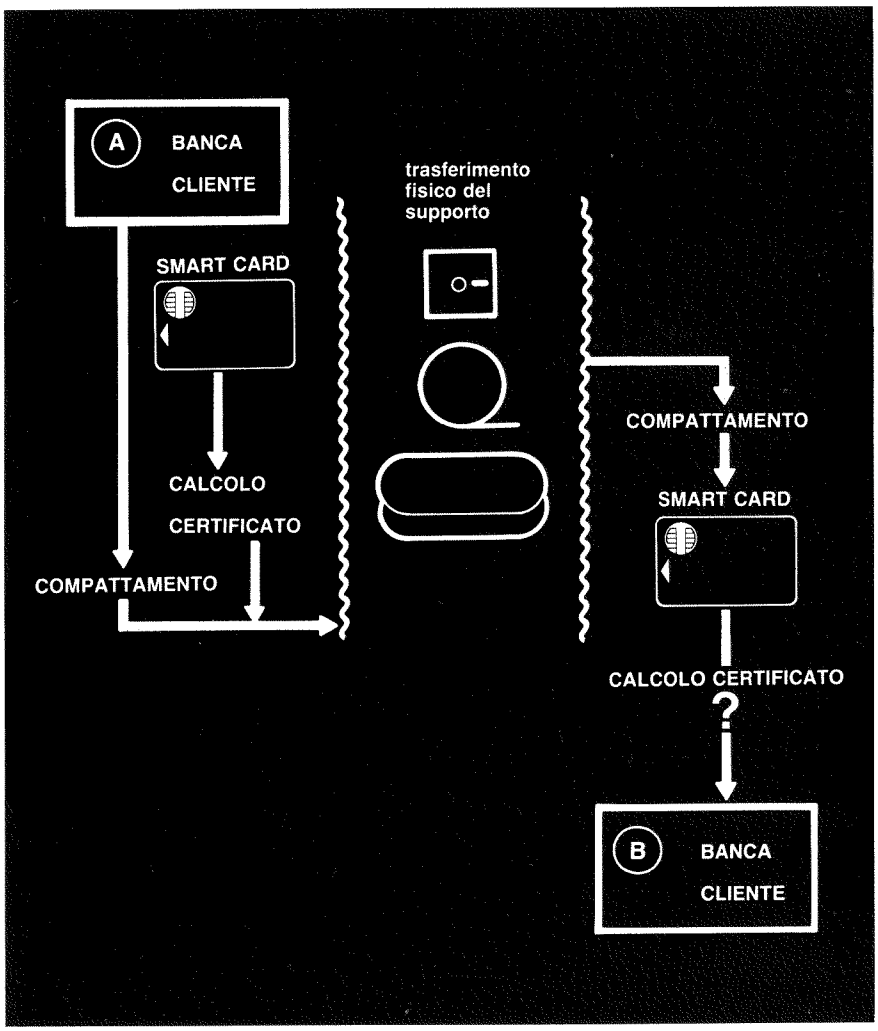


Fig. 6 - Protezione del messaggio mediante crittografia (A = emittente, B = ricevente, C = il certificato viene controllato con l'equivalente registrato in coda al supporto).



Con la *smart card* si può utilizzare una chiave segreta che viene generata dalla carta stessa partendo dai dati contenuti in memoria, protetti ed innaccessibili, e dalla chiave banale emessa dal sistema.

Anche in questo caso è necessaria una carta, all'arrivo del messaggio, in grado di svolgere lo stesso procedimento in senso inverso.

Lo schema operativo è illustrato nella figura 6.

Per quanto riguarda l'applicazione della *smart card* CP8 per il controllo degli *accessi fisici,* valgono ovviamente le stesse considerazioni fin qui svolte relativamente alle caratteristiche di sicurezza della carta e del sistema di emissione.

L'utilizzazione del CP8 come chiave d'accesso a un locale avviene attraverso un lettore generalmente "off-line", in quanto in grado di controllare automaticamente le informazioni contenute nella carta, installato in prossimità dell'accesso che si vuole proteggere, di dimensioni contenute.

Questo lettore è dotato di un meccanismo in grado di far scattare una serratura e/o accendere una spia luminosa, per indicare se la carta presentata è stata accettata o rifiutata.

Come per il controllo degli accessi logici, l'accettazione della carta è subordinata alla presentazione del codice personale corretto.

È inoltre possibile stabilire differenti livelli di abilitazione, codificandoli opportunamente nella memoria della carta, selezionando quindi i diritti di accesso e ottenendo una maggiore flessibilità dell'intero sistema.

4. Conclusioni

La *smart card,* resa possibile dalle moderne tecnologie elettroniche, costituisce un vero e proprio salto in avanti dei sistemi di sicurezza logica e fisica.

Per le sue caratteristiche intrinseche, sia funzionali che economiche, la *smart card* rappresenta una risposta efficace ed efficiente ad una problematica presente ormai nei più svariati settori della società in cui viviamo.

CAI per l'handicap: un supporto alla formazione professionale

MARGHERITA CANTARELLA

Bull Italia
Sviluppo CAI
Milano

1. Introduzione

Il Progetto N.I.T. (Nuove Tecnologie Informatiche) è nato dal desiderio del CIFAP (Consorzio per l'Istruzione e la Formazione Artigiana e Professionale) di sperimentare nuove tecnologie nella formazione professionale e dalla capacità del gruppo CAI della Bull Italia di venire incontro a tale desiderio.

Diversi enti nazionali e comunitari[1] hanno approvato e sostengono questo progetto pilota il cui obiettivo scientifico è verificare se l'introduzione di prodotti CAI nel programma di formazione professionale di giovani portatori di handicap costituisce per loro un valido aiuto nel processo di apprendimento.

La metodologia e gli strumenti utilizzati e le attività svolte per raggiungere l'obiettivo sono descritti in questo articolo.

2. Definizione del contesto

Il CIFAP ha individuato, a seguito di indagini nell'ambito del territorio su cui opera, cinque professioni di livello medio-basso che offrono garanzie di "sopravvivenza" nei prossimi anni, non sono cioè, soggette a diventare rapidamente obsolete.

A queste professioni dovevano essere avviati 45 giovani con handicap fisici e psichici, di età compresa tra i 18 e i 25 anni.

La metodologia utilizzata, come per precedenti esperienze, prevedeva:

- l'attività di una equipe specializzata per la formazione di base dei soggetti;
- la collaborazione di varie aziende per lo svolgimento del tirocinio professionale.

La novità, nel progetto N.I.T., è rappresentata dall'aggiungersi di un nuovo strumento di formazione: il computer.

Esperti CAI della Bull Italia sono stati, quindi, impegnati nella realizzazione di cinque courseware, ciascuno relativo ad una delle professioni individuate e cioè: ausiliario, operaio di linea, saldatore, impiegato amministrativo e contabile.

I package CAI rappresentano un supporto al tirocinio professionale nel senso che:

- verificano il possesso delle capacità di base richieste per affrontare con successo la professione;
- forniscono, attraverso il feedback, un rinforzo e un approfondimento.

3. Fasi del progetto

Il progetto, partito a fine '85 e ormai concluso per quanto riguarda la produzione del courseware, ha avuto inizio con una fase di formazione per il personale CIFAP.

Trattandosi di persone con nessuna esperienza d'informatica e, in particolare, di CAI, gli esperti Bull hanno proposto un programma personalizzato che comprendeva:

- una parte introduttiva di familiarizzazione all'uso del PC;
- un corso di metodologia e tecniche CAI;
- un periodo di on-the-job-training durante il quale i partecipanti hanno realizzato una breve unità didattica CAI su un argomento scelto da loro.

Terminata la formazione, è stato costituito il team di progetto per la realizzazione del primo courseware.

Le figure professionali coinvolte, per il primo come per tutti gli altri package CAI, sono state: l'esperto della materia (rappresentante di un'azienda o del sindacato), l'esperto di problematiche dell'handicap (CIFAP), l'esperto CAI (Bull).

Il personale CIFAP ha svolto una serie di interviste a lavoratori che svolgono le attività per le quali andava sviluppato il courseware; i risultati sono stati elaborati dal team per individuare le capacità di base necessarie per poter svolgere il lavoro in questione.

Tali capacità costituiscono il contenuto dei package CAI e sono state organizzate in strutture gerarchiche,

(1) Comunità Economica Europea, Consiglio d'Europa, Ministero del Lavoro, Ministero della Ricerca Scientifica, Regione Lombardia, Provincia di Milano, Comuni di Sesto S. Giovanni, Cologno e Cinisello Balsamo.

scomponendo le più complesse fino a determinare le capacità elementari (Fig. 1).

Per ogni capacità sono state progettate delle domande sotto forma di situazioni di fronte alle quali lo studente deve prendere una decisione.

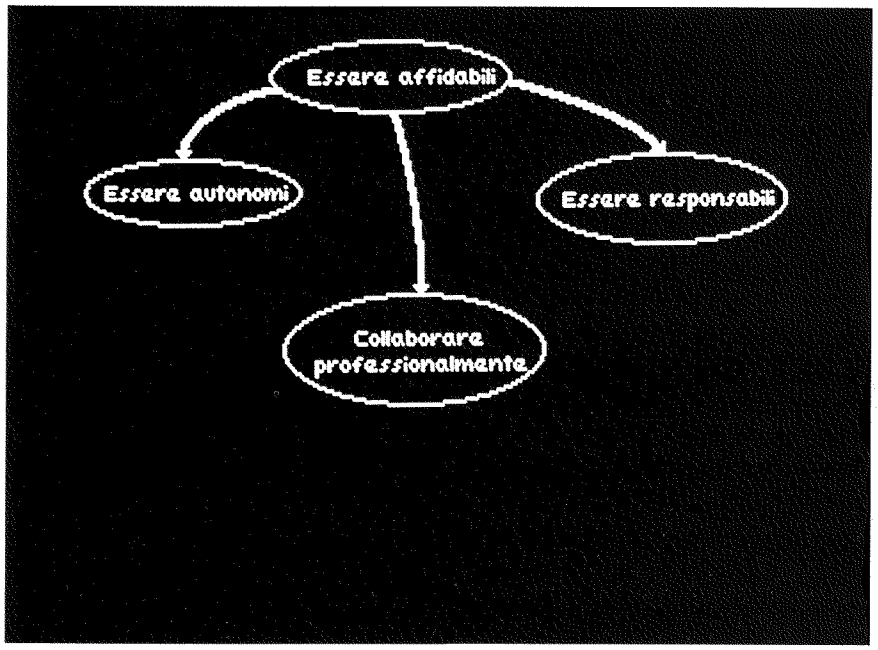
Le situazioni si riferiscono sia al contesto familiare e quotidiano (Fig. 2.a) sia a quello di lavoro (Fig. 2.b).

Gli esperti CIFAP hanno scelto di ridurre al minimo le parti tutoriali, utilizzando quasi esclusivamente il feedback alle risposte come strumento di rinforzo e approfondimento.

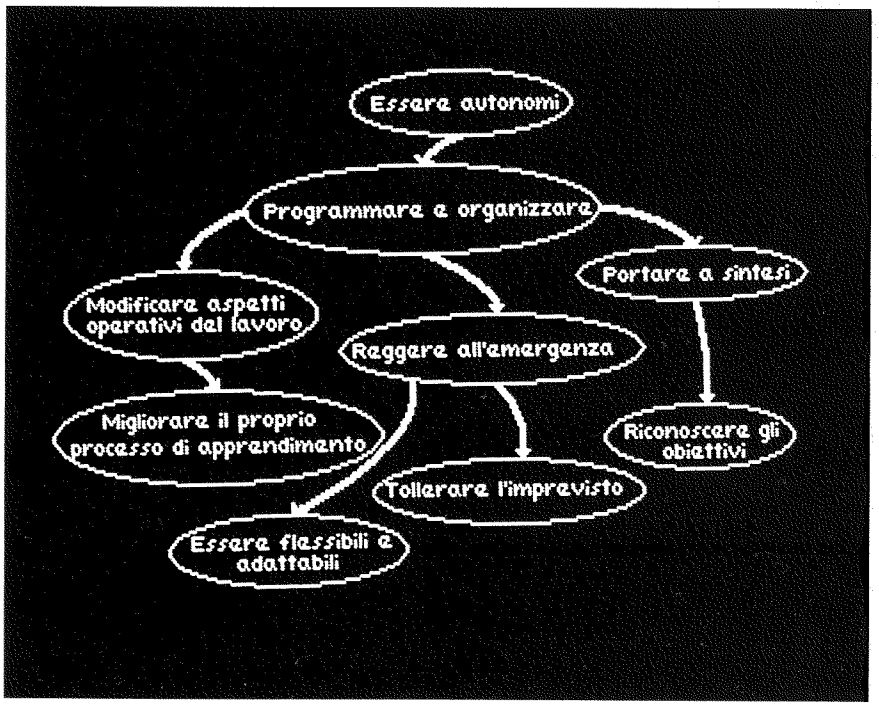
Parallelamente alla progettazione dei courseware, gli esperti Bull hanno realizzato una procedura per l'acquisizione ed elaborazione dei dati relativi alle performance degli studenti e per la gestione dell'aula elettronica. (Maggiori dettagli su tale procedura sono forniti in un successivo paragrafo).

Realizzati i courseware, gli utenti hanno cominciato ad utilizzarli nello stesso periodo in cui svolgevano il tirocinio in azienda.

Le sessioni in autoistruzione si svolgono ad intervalli stabiliti dagli esperti CIFAP in base alle caratteri-



a)

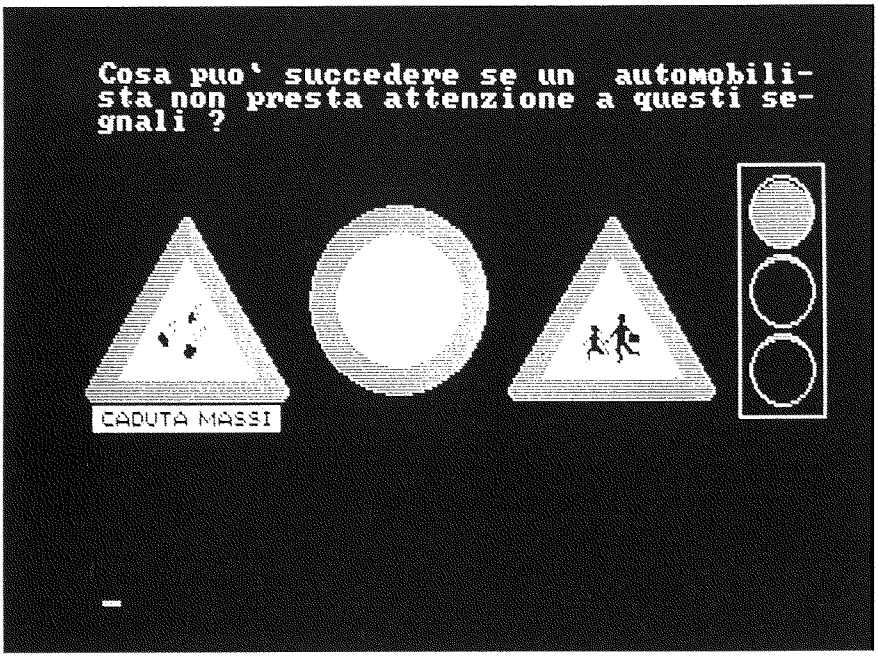


b)

Fig. 1 - Parte della struttura gerarchica dei contenuti del package "Impiegato amministrativo": a) capacità che concorrono a costituire "Essere affidabili", b) capacità che concorrono a costituire "Essere autonomi".

Fig. 2 - Per ogni situazione sono state progettate delle domande sotto forma di situazioni in cui l'allievo deve prendere una decisione. Le situazioni si riferiscono sia al contesto familiare e quotidiano (fig. a) sia a quello di lavoro (fig. b).

a)



b)

stiche e alle esigenze dei singoli utenti.

La molteplicità di fruizioni dei package consente di rilevare, attraverso i dati memorizzati, progressi, regressi, stazionarietà dei soggetti sulle aree del courseware (equivalenti ad altrettante capacità).

4. I Courseware

Le caratteristiche delle professioni individuate hanno, naturalmente, influito sulla struttura e la complessità dei singoli courseware, in quanto sono diverse le capacità richieste.

AUSILIARIO: è il profilo meno complesso; la capacità più rilevante è un buon livello di autonomia esecutiva nello svolgimento delle operazioni richieste dalla mansione.

Le situazioni proposte, quindi, riguardano il contesto lavorativo e mirano ad evidenziare che anche attività semplici richiedono precisione, organizzazione e senso di responsabilità (Fig. 3).

Dal punto di vista della struttura, il package prevede una diversificazione di percorso, dopo una parte comune, in quanto alcune prestazioni sono richieste solo in enti specifici.

La qualifica di ausiliario, infatti, può indicare il bidello delle scuole, l'addetto alle pulizie degli ospedali, l'addetto alla distribuzione dei pasti nelle mense, ecc.

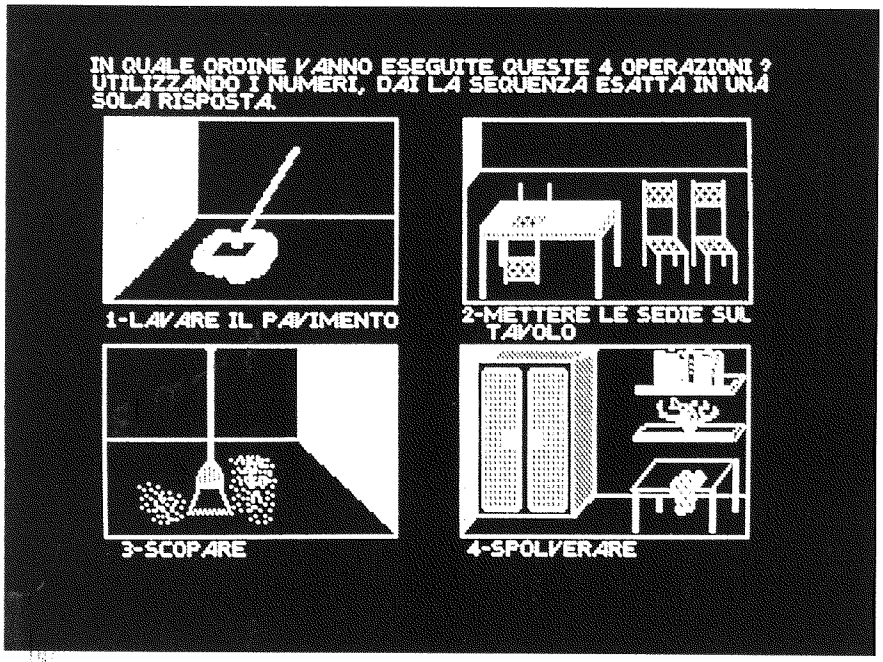


Fig. 3 - Una delle situazioni proposte per la professione di ausiliario (capacità di organizzare il lavoro).

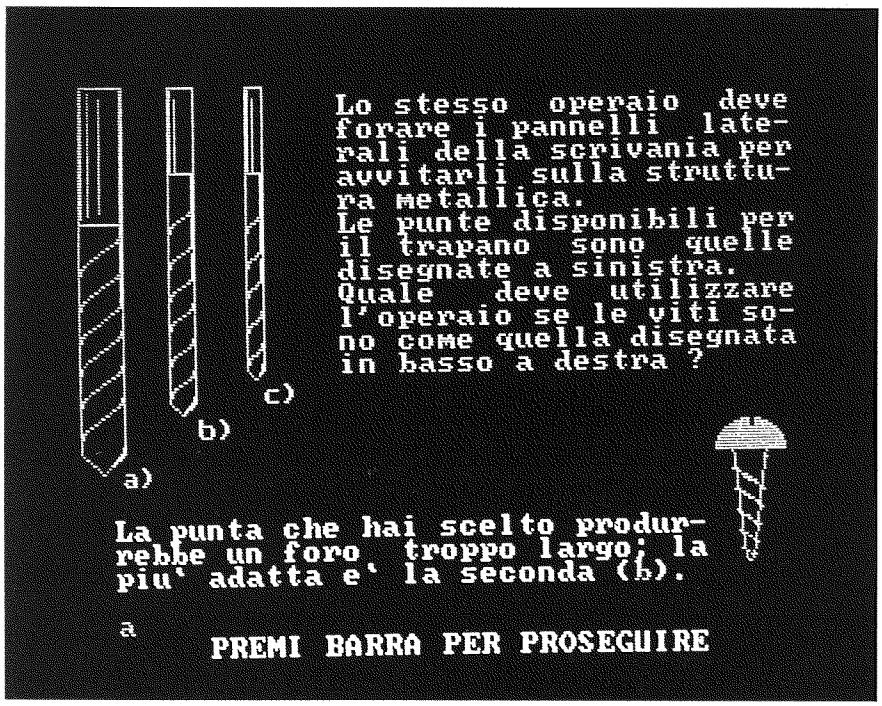


Fig. 4 - Una delle situazioni proposte per l'operaio di linea (capacità di selezionare).

OPERAIO DI LINEA (Fig. 4): È un'attività che richiede un buon livello di coordinazione, l'assoluto rispetto delle istruzioni ricevute, nessun tipo di iniziativa personale e, di conseguenza capacità di tollerare lo stress.

Le situazioni proposte, quindi, insistono sulla caratteristica di attività strettamente esecutiva; per "ammorbidire" la sensazione di lavoro meccanico, uno spazio è riservato a valorizzare il ruolo dell'operaio di linea all'interno del ciclo produttivo.

SALDATORE (Fig. 5): Nello svolgimento del suo lavoro, il saldatore ha maggiore autonomia rispetto all'operaio di linea.

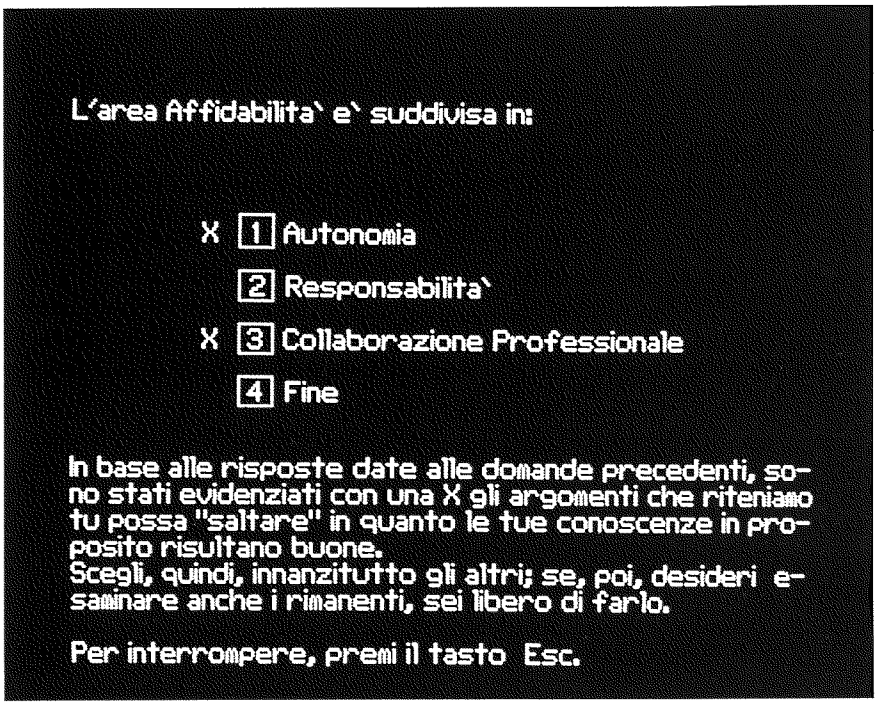
Non gli vengono impartite istruzioni dettagliate, per cui deve essere in grado di integrare le disposizioni ricevute con la propria esperienza e di prendere decisioni operative, pur nei limiti delle sue competenze. Lavora spesso in gruppo, quindi deve possedere buone capacità di relazione e interazione all'interno del gruppo stesso.

Le situazioni proposte riguardano in parte la vita quotidiana e in parte il contesto lavorativo per dimostrare che le esperienze acquisite altrove possono rivelarsi utili anche nello svolgimento del lavoro.

IMPIEGATO AMMINISTRATIVO: Entrano in gioco, in questo

Fig. 5 - Una delle situazioni proposte per il saldatore (comprendere le istruzioni).

Fig. 6 - Profilo impiegato amministrativo.



profilo professionale, componenti quali l'affidabilità, la motivazione, la disponibilità al cambiamento di skill, la corretta valutazione delle risorse mentali e psicologiche da impiegare per raggiungere gli obiettivi.

Poichè molteplici capacità concorrono a definire tali componenti, il courseware risulta piuttosto articolato; sono stati individuati, perciò, alcuni punti in cui effettuare una verifica tramite test. In base ai risultati dello studente, il percorso didattico si differenzia; il courseware segnala all'allievo quali capacità deve esercitare e quali può considerare acquisite. Lo studente può selezionare sia gli argomenti consigliati, sia le parti relative alle capacità di cui risulta già in possesso; tuttavia, una nota gli suggerisce di dare priorità agli argomenti nei quali risulta meno "capace" (Fig. 6).

CONTABILE (Fig. 7): Le caratteristiche principali di questo profilo professionale sono precisione, metodicità, organizzazione.

Il package, quindi, propone situazioni relative alle capacità di selezione e classificazione, alla corretta gestione dei ritmi di lavoro, alla relazione tra competenza e affidabilità.

Inoltre, poiché l'attività del contabile è spesso ripetitiva, una parte del

Fig. 7 - Una delle situazioni proposte per il contabile (capacità di classificare).

courseware è dedicata alla motivazione, per sottolineare che l'esistenza di questa componente è importante per affrontare gli aspetti meno gradevoli del lavoro.

5. Ambiente di fruizione dei courseware.

Date le caratteristiche degli allievi e le esigenze di analisi dei risultati, è stato realizzato un ambiente integrato per la fruizione dei courseware.

Si tratta di un'aula elettronica costituita da 15 Personal Computer collegati tra loro in rete locale: un PC ha funzionalità di *tutor* e di *server* di rete, gli altri sono stazioni di lavoro per gli studenti.

I componenti software dell'aula sono la Local Area Network, un Monitor, i Courseware e un'Interfaccia.

Le funzionalità centralizzate offerte da questo ambiente sono:

- Archivio Courseware
- Controllo
- Raccolta dati
- Invio messaggi

Il docente, quindi, può (dal PC *server*):

- Distribuire i Courseware agli studenti
- Verificare lo stato delle stazioni di lavoro
- Controllare le attività di uno o di tutti i posti studente
- Inviare messaggi

Durante la fruizione dei Courseware i dati relativi ai vari studenti sono registrati sul PC *server* per successive elaborazioni a fini statistici e di analisi.

6. Conclusioni

Anche se il progetto N.I.T. può considerarsi concluso per ciò che concerne la produzione dei courseware, la realizzazione delle procedure per la gestione dell'aula e l'installazione dell'aula stessa, non è possibile presentare dati consuntivi completi.

La fruizione dei package CAI, infatti, avviene durante il periodo di tirocinio degli utenti e non tutti i tirocini sono partiti alla stessa data: di conseguenza non tutti sono conclusi.

È possibile, comunque, fare alcune considerazioni.

Innanzitutto, lo strumento computer è stato generalmente accettato senza problemi.

Inoltre, i dati finora rilevati dagli esperti CIFAP indicano che i soggetti che superano con successo le problematiche dei Courseware raggiungono anche una buona integrazione all'interno delle aziende, assumendo un ruolo produttivo.

Infatti, l'80% di coloro che hanno già concluso il tirocinio ha ottenuto l'assunzione.

I risultati definitivi saranno pubblicati al termine dell'esperienza, a cura degli specialisti CIFAP.

Un anniversario storico: la teoria dell'informazione

WILLIAM ASPRAY

Charles Babbage Institute
Minneapolis (USA)

Quarant'anni fa Claude Shannon pubblicava l'articolo "The mathematical theory of communication", *una pietra miliare nella storia della scienza dell'informazione.*

La maturazione culturale che portò a tale risultato è tratteggiata in questo scritto, tratto dall'articolo "The scientific conceptualization of information", *pubblicato dall'Autore in "Annals of the history of computing".*

La teoria matematica della comunicazione

Negli anni 40, mentre lavorava ai Bell Laboratories sui problemi relativi alle trasmissioni radio e telegrafiche, Shannon sviluppò una teoria generale della comunicazione riguardante la trasmissione di ogni sorta di informazione da un punto ad un altro nello spazio o nel tempo. Il suo scopo era quello di fornire definizioni tecniche specifiche di concetti sufficientemente generali da applicare in ogni situazione in cui l'informazione venga manipolata o trasmessa, ossia concetti come informazione, disturbo, trasmittente, segnale, ricevente e messaggio.

Alla base della teoria c'era una nuova concezione della informazione.

Per rendere la teoria della comunicazione una disciplina scientifica, a Shannon occorreva una definizione precisa dell'informazione, che trasformasse quest'ultima in un parametro fisico quantificabile. Egli realizzò questa trasformazione distinguendo l'informazione dal suo significato. Riservò il termine "significato" al contenuto intrinseco di un particolare messaggio. Utilizzò invece il termine "informazione" per riferirsi al numero dei diversi messaggi che potrebbero essere trasportati su un canale in funzione della lunghezza del messaggio e del numero di scelte dei simboli usati in ogni punto del messaggio.

Secondo Shannon, l'informazione è una misura di regolarità (contrapposta alla casualità) in quanto essa indica il numero di messaggi possibili tra cui viene scelto un particolare messaggio da inviare. Quanto maggiore il numero di possibilità, tanto maggiore la quantità di informazioni trasmesse, poiché il messaggio effettivo si differenzia rispetto ad un numero di possibilità alternative.

Shannon riconobbe l'importanza che i lavori precedenti sull'ingegneria della comunicazione, in particolare il lavoro di Harry Nyquist e R.V. Hartley, rivestivano per la teoria generale dell'informazione.

Il recente sviluppo di vari metodi di modulazione come PGM e PPM ha intensificato l'interesse per una teoria generale della comunicazione. La base di una tale teoria è contenuta nelle importanti memorie di Nyquist e Hartley. In questo articolo estenderemo la teoria includendo numerosi nuovi elementi (Rif. [3], p. 31-32).

Nyquist stava conducendo le sue ricerche sul problema dell'incremento della velocità di trasmissione su linee telegrafiche ai Bell Laboratories, quando pubblicò uno studio sulla trasmissione dell'informazione (1924).

Questo studio riguardava due fattori che impattano sulla massima velocità con cui l'informazione può essere trasmessa per telegrafo: forma dei segnali e scelta dei codici. Nyquist scriveva:

Il primo fattore riguarda la miglior forma da imprimere sul mezzo di trasmissione, in modo da garantire la più elevata velocità senza indebite interferenze nel circuito considerato o in quelli adiacenti, mentre il secondo fattore tratta la scelta di codici che permettono di trasmettere la massima quantità possibile di informazione con un dato numero di elementi di segnale (Rif. [2], p. 324).

Lo studio di Nyquist presentava una sezione teorica di grande importanza, intitolata "Possibilità teoriche consentite dall'utilizzo di codici con numeri diversi di valori correnti".

In questa sezione Nyquist introdusse la relazione logaritmica che governa la trasmissione di informazioni.

Nyquist dimostrò che la velocità con cui l'informazione può essere trasmessa su un circuito telegrafico obbedisce all'equazione

$$W = k \log m$$

dove W è la velocità di trasmissione dell'informazione, m è il numero di valori corretti che possono essere trasmessi e k è una costante.

Creò inoltre la tabella seguente, con cui illustrò i vantaggi dell'utilizzo di un numero più alto di valori correnti per la trasmissione di messaggi.

Numero di valori correnti	Quantità relativa di informazione che può essere trasmessa con un dato numero di elementi di segnale
2	100
3	158
4	200
5	230
8	300
16	400

Sebbene lo studio di Nyquist fosse essenzialmente empirico e riguardasse problemi d'ingegneria - e sebbene egli usasse il termine *intelligence,* che mascherava la differenza fra informazione e significato - il suo lavoro fu di notevole importanza in quanto formulava per la prima volta una legge logaritmica per la comunicazione e uno studio dei limiti teorici dei codici ideali per la trasmissione di informazioni. Shannon fornì in seguito una regola logaritmica più generale come legge fondamentale della teoria della comunicazione, che stabiliva che la quantità di informazione è direttamente proporzionale al logaritmo del numero dei messaggi possibili. La legge di Nyquist diventò un caso specifico della teoria di Shannon, in quanto il numero di valori è direttamente correlato al numero dei simboli che possono essere trasmessi. Nyquist era consapevole di questa relazione, come indica la sua definizione di velocità di trasmissione.

Per velocità di trasmissione di informazione s'intende il numero di caratteri, rappresentanti differenti lettere, figure, ecc., che può essere trasmesso in un dato intervallo di tempo. (Rif. [2], p. 333).

Con "lettere, figure, ecc." egli intendeva una misura proporzionale all'entità che Shannon chiamerà più tardi "bits d'informazione". La tabella di Nyquist, che riporta la quantità *relativa* di "intelligence" trasmessa, illustra il guadagno in termini di informazione conseguente ad un numero più elevato di scelte possibili. Il fatto che egli elenchi la quantità *relativa,* dimostra che egli si rese conto dell'importante relazione esistente tra numero di valori e quantità di informazione da trasmettere.

Questa relazione è alla base della teoria della comunicazione di Shannon. Tuttavia Nyquist non generalizzò il suo concetto al di fuori delle trasmissioni telegrafiche.

Anche l'altro predecessore di Shannon nella teoria dell'informazione, R. Hartley, era un ricercatore dei Bell Laboratories.

Hartley si proponeva di stabilire una misura quantitativa per confrontare la capacità dei vari sistemi di trasmettere informazioni. Egli sperava di fornire una teoria abbastanza generale che includesse telegrafia, telefonia, televisione e trasmissioni di immagini sia su filo che via radio. Il suo studio (1928) cominciò con un tentativo di stabilire limiti teorici di trasmissione dell'informazione in situazioni ideali. Questo importante passo lo allontanò dagli studi empirici adottati dalla maggior parte dei ricercatori precedenti e lo condusse verso una teoria matematica della comunicazione.

Prima di rivolgersi ai problemi concreti d'ingegneria, Hartley si dedicò a "considerazioni più astratte". Iniziò cercando anzitutto di distinguere una nozione di informazione riconducibile all'utilizzo in un contesto scientifico. Egli si rese conto che ogni definizione di "informazione" scientificamente utilizzabile deve essere basata su quelle che egli chiamò considerazioni "fisiche" invece che "psicologiche". Intendeva dire che l'informazione è un'idea che riguarda grandezze fisiche e non deve essere confusa con il significato del messaggio.

La capacità di un sistema di trasmettere una particolare sequenza di simboli dipende dalla possibilità di distinguere all'estremità ricevente i risultati delle varie selezioni effettuate all'estremità trasmittente. L'operazione di riconoscimento, dal record ricevuto, della sequenza di simboli selezionata all'estremità trasmittente, può essere realizzata anche da chi non conosca l'alfabeto Morse.
Potremmo fare ciò altrettanto bene per una sequenza che rappresenti un messaggio scelto consapevolmente e per un'altra sequenza inviata dal dispositivo di selezione automatica già citato. Tuttavia un operatore esperto direbbe che la sequenza inviata dal dispositivo automatico non è comprensibile. Ciò è dovuto al fatto che solo ad un limitato numero di sequenze possibili sono stati assegnati significati comuni a lui e all'operatore trasmittente. Pertanto il numero di simboli disponibile all'operatore trasmittente per certe sue selezioni è limitato in questo caso da considerazioni psicologiche piuttosto che fisiche.
Altri operatori che utilizzano altri codici potrebbero effettuare altre selezioni.
Perciò, nella valutazione della capacità del sistema fisico di trasmettere informazioni, dobbiamo ignorare il problema dell'interpretazione, rendere ogni selezione perfettamente arbitraria e basare i nostri risultati sulla possibilità del ricevente di distinguere il risultato della selezione di un simbolo rispetto alla selezione di un altro. In questo modo infatti i fattori psicologici e le relative variazioni sono eliminati e diventa possibile fissare una misura quantitativa dell'informazione basata esclusivamente su considerazioni fisiche. (Rif. [1], p. 537-538).

Pertanto Hartley distinse le considerazioni fisiche da quelle psicologiche, cioè il significato dall'informazione. Quest'ultima la definì come il numero dei messaggi possibili, indipendentemente dalla ricchezza di significato. Egli utilizzò questa definizione di informazione per formulare una legge logaritmica per la trasmissione di informazioni:

$$H = K \log s^n$$

dove H è la quantità di informazione, K è una costante, n è il numero dei simboli presenti nel messaggio, s è l'ampiezza dell'insieme di simboli e perciò s^n è il numero delle sequenze simboliche possibili con lunghezza n specificata. Questa legge riguarda anche il caso della telegrafia e include la precedente legge di Nyquist.

Una volta stabilita la legge per la trasmissione "discreta" di informazioni, Hartley dimostrò come può essere modificata per gestire la trasmissione "continua" d'informazioni, come nel caso della trasmissione fonica.

Successivamente Hartley si dedicò al problema delle interferenze.

Egli si interessò specialmente alle interferenze dovute all'accumulo ed alla successiva scarica di energia per via induttiva e capacitiva, fonti di disturbi che preoccupavano notevolmente gli ingegneri elettrotecnici di quel tempo.

Hartley era giunto a molte delle più importanti idee della teoria matematica della comunicazione: la differenza tra informazione e significato, l'informazione come quantità fisica, la legge logaritmica per la trasmissione di informazioni e il concetto di disturbo come impedimento nella trasmissione di informazioni.

Lo scopo di Hartley era stato quello di costruire una teoria in grado di valutare le informazioni trasmesse

con una qualsiasi delle tecnologie standard di comunicazione.

Partendo da queste idee, Shannon sviluppò una teoria generale della comunicazione, non limitata allo studio di tecnologie previste specificatamente per la comunicazione. Poco dopo, un collaboratore di Shannon, Warren Weaver, descrisse chiaramente la teoria.

La parola "comunicazione" verrà qui usata in un senso molto ampio, per includere tutte le procedure con cui una mente può influenzare un'altra. Naturalmente ciò riguarda non solo la parola scritta e orale, ma anche musica, pittura, teatro, balletto e, in definitiva, l'insieme dei comportamenti umani. Sotto certi aspetti può essere desiderabile utilizzare una definizione di comunicazione ancora più ampia, cioè una definizione che includa procedure per mezzo delle quali un meccanismo (ad esempio un'apparecchiatura automatica per seguire un aereo e calcolare le sue probabili posizioni future) può influenzare un altro meccanismo (ad esempio un missile guidato contro quell'aereo). (Rif. [4], Introductory Note).

Quello che all'inizio era stato solo uno studio della trasmissione su linee telegrafiche, fu poi sviluppato da Shannon in una teoria generale della comunicazione, applicabile a telegrafo, telefono, radio, televisione, macchine da calcolo e, in definitiva, a qualsiasi sistema, fisico o biologico, in cui l'informazione venga trasferita o manipolata nel tempo o nello spazio.

Nella teoria di Shannon un sistema di comunicazione consiste di cinque componenti, come illustrato nella Figura 1. (Rif. [4], p. 34). Questi componenti sono:

1. *Una fonte di informazioni che produce un messaggio o una sequenza di messaggi da trasmettere al terminale ricevente...*
2. *Un trasmettitore che opera sul messaggio in modo tale da produrre un segnale adatto alla trasmissione sul canale...*
3. *Il canale è soltanto il mezzo usato per trasmettere il segnale dal trasmettitore al ricevitore...*
4. *Il ricevitore esegue normalmente l'operazione inversa rispetto a quella del trasmettitore, ricostruendo il messaggio dal segnale...*
5. *La destinazione è la persona (o la cosa) per cui è inteso il messaggio.*

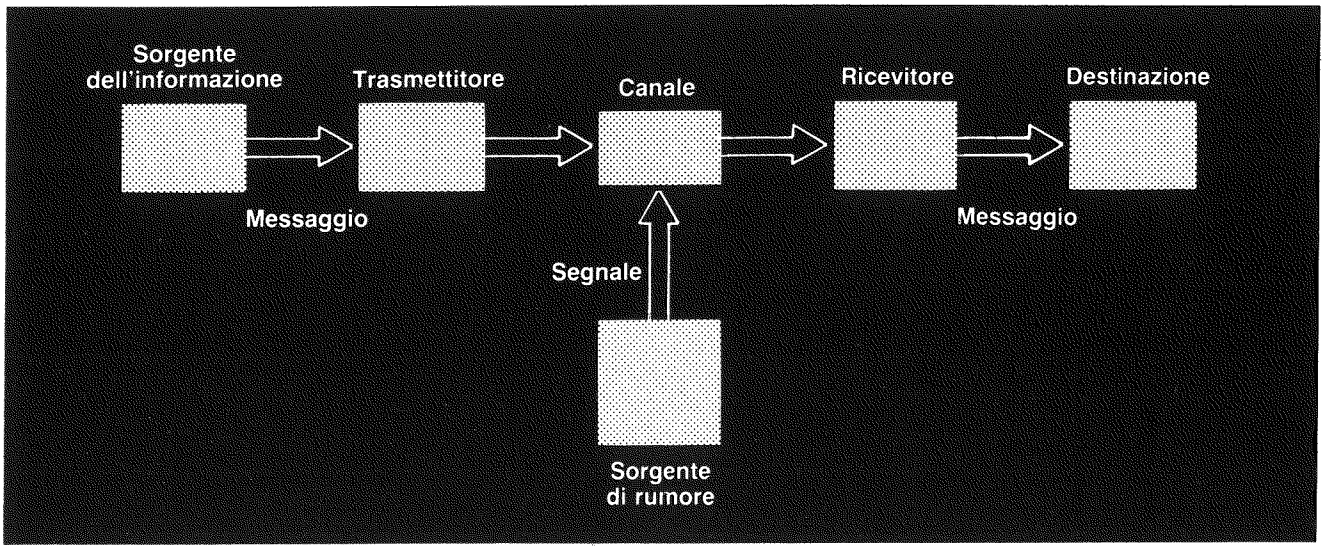
L'importanza di questa caratterizzazione è la sua applicabilità a una vasta gamma di problemi di comunicazione, purchè i cinque componenti siano interpretati nel modo appropriato. Per esempio, si applica bene anche a conversazioni fra individui, interazioni fra macchine e anche alla comunicazione fra parti di un organismo. Se propriamente interpretate, le comunicazioni tra lo stomaco e il cervello, o tra l'obiettivo e il missile guidato, possono essere considerate come esempi di un sistema di comunicazione. Utilizzando la teoria di Shannon, le connessioni tra mondo biologico e mondo fisico, prima sconosciute, possono essere evidenziate.

Mentre Hartley riconobbe che doveva essere tracciata una distinzione tra informazione e significato, Shannon affinò tale distinzione fornendo la prima definizione di informazione sufficientemente precisa per un discorso scientifico. Weaver dimostrò l'importanza di questa definizione.

La parola "informazione", in questa teoria, è usata in un senso speciale che non deve essere confuso con il normale utilizzo. In particolare, informazione non deve essere confusa con significato.
Infatti, due messaggi uno dei quali è fortemente carico di significato e l'altro è puramente insulso, dal punto di vista dell'informazione possono essere esattamente equivalenti.
In realtà, la parola informazione nella teoria della comunicazione si riferisce non tanto a quello che si dice, quanto a ciò che si potrebbe dire. Cioè, l'informazione è la misura della libertà individuale di scelta nella selezione di un messaggio. Se ci si trova di fronte ad una situazione molto elementare in cui bisogna scegliere fra due messaggi alternativi, viene arbitrariamente detto che l'informazione, associata a questa situazione, è unitaria. Può trarre in inganno (sebbene sia spesso conveniente) dire che l'uno e l'altro dei messaggi trasporta informazioni unitarie. Il concetto di informazione non si applica ai messaggi individuali (come potrebbe essere per il concetto di significato), ma piuttosto alla situazione nel suo complesso, poichè l'informazione unitaria indica che in questa situazione si ha una certa libertà di scelta, nella selezione di un messaggio, che è conveniente considerare come una quantità standard o unitaria. (Rif. [4], p. 8-9).

Shannon trovò che l'informazione può essere misurata con una funzione monotonica crescente, purchè il numero dei possibili messaggi sia finito. Egli scelse fra esse la funzione logaritmica per lo stesso motivo di Hartley: cioè che si accorda bene

Fig. 1 - Schematizzazione di un sistema generale di comunicazione (Shannon e Weaver, 1949, Rif. [4]).



con la nostra intuizione. Noi intuiamo che due schede perforate portano il doppio di informazioni rispetto ad una sola. Supponendo che una scheda possa contenere n simboli, due schede conterranno n^2 combinazioni. La funzione logaritmica misura una capacità delle due schede pari a $\log n^2 = 2 \log n$ bit di informazione, il doppio di quanto contenuto in una sola scheda ($\log n$). Shannon scelse la base 2 per questa misura logaritmica, dato che $\log_2$ assegna un'unità di informazione ad uno switch a due posizioni.

Quindi N switches a due posizioni possono memorizzare $\log_2 2^N (=N)$ cifre binarie di informazione. Se ci fossero N scelte equiprobabili, la quantità di informazioni verrebbe data da $\log_2 N$. Shannon generalizzò questa equazione per la situazione di non-equiprobabilità, in cui la quantità di informazione H risulta data da

$$H = -(p_1 \log_2 p_1 + ... + p_n \log_2 p_n)$$

se le scelte hanno probabilità $p_1, ..., p_n$.

Shannon trovò che questa formulazione di informazione è strettamente legata al concetto di *entropia*, dato che il parametro informazione misura la regolarità del canale di comunicazione.

Le quantità sotto forma di $H = -P_i \log P_i$ (la costante K corrisponde semplicemente alla scelta di un'unità di misura) svolgono un ruolo centrale... come misure di informazione, scelta ed incertezza.
La forma di H corrisponde a quella di entropia come viene definita in certe formule di meccanica statistica, in cui P_i è la probabilità di un sistema di essere nella cella i del suo spazio di fase. Quindi H è, per esempio, l'H del famoso teorema H di Boltzmann. (Rif. [4]).

L'entropia ha una lunga storia nella fisica e durante il ventesimo secolo viene strettamente associata con la quantità di informazioni in un sistema fisico.

Scrive Weaver:

Il lavoro del Dr. Shannon si rifà, come aveva evidenziato von Neumann, allo studio di Boltzmann sulla fisica statistica (1894), nel senso che l'entropia è legata all'"informazione mancante", in quanto si riferisce al numero di alternative possibili rimaste a disposizione di un sistema fisico, dopo che tutte le relative informazioni microscopiche siano state registrate. Szilard estese quest'idea ad un discorso generale sull'informazione nella fisica, e von Neumann trattò le informazioni nella meccanica quantistica e nelle fisica delle particelle. (Rif. [4], p. 3).

Questa stretta relazione tra informazione ed entropia non è sorprendente, in quanto l'informazione è correlata al grado di libertà di scelta disponibile durante la costruzione di messaggi. Questo legame fra termodinamica, meccanica statistica e teoria della comunicazione suggerisce che quest'ultima implica una proprietà basilare dell'universo fisico e non è semplicemente un sottoprodotto scientifico della moderna tecnologia della comunicazione.

Shannon utilizzò la sua teoria per dimostrare diversi risultati teorici relativi ai sistemi di comunicazione e per proporre applicazioni all'industria delle telecomunicazioni.

Egli stabilì dei limiti teorici applicabili a sistemi pratici di comunicazione. La sua teoria fornì una nuova definizione di *informazione* applicabile ad un'ampia varietà di ambienti fisici e mise a disposizione un approccio matematico allo studio dei problemi della trasmissione e del trattamento dell'informazione. Shannon e Weaver continuarono lo studio teorico di questo argomento. Nel frattempo, la teoria costituì la base di studi interdisciplinari sull'informazione applicati ai sistemi di elaborazione elettronica e a sistemi vari di natura fisica e biologica.

Riferimenti bibliografici

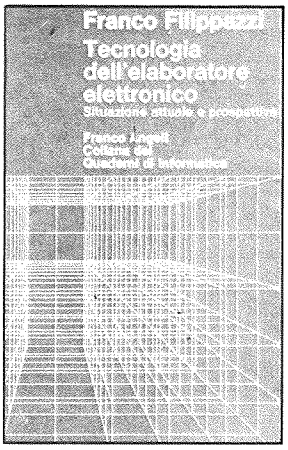
[1] HARTLEY, R.V. 1928. Transmission of information. Bell System Technical Journal 7, 535-553.

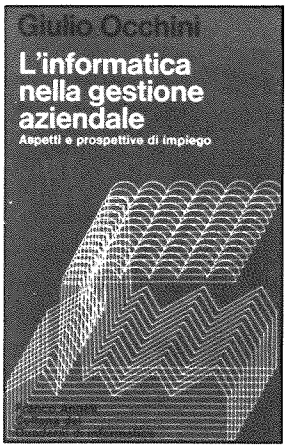
[2] NYQUIST, H. 1924. Certain factors affecting telegraph speed. Bell System Technical Journal 3, 324-346.

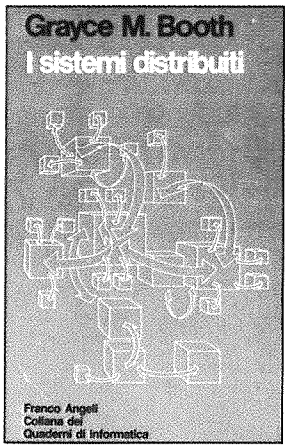
[3] SHANNON, C. 1948. The mathematical theory of communication. Bell System Technical Journal 27, 3 and 4, 379-423 and 623-656.

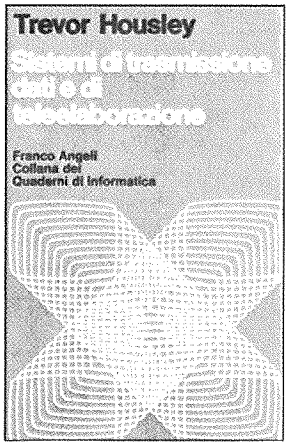
[4] SHANNON, C., and W. WEAVER. 1949. The Mathematical Theory of Communication. Urbana, University of Illinois Press.

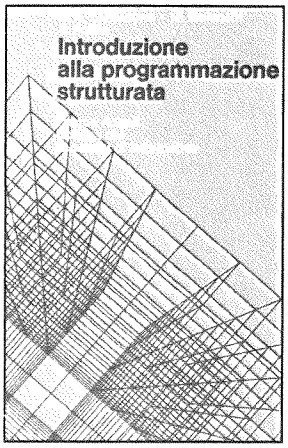
I TESTI DEI
QUADERNI DI INFORMATICA

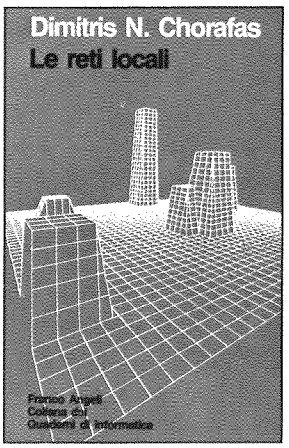




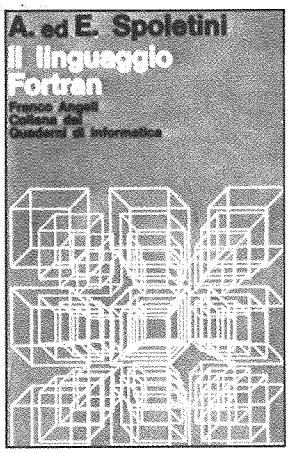


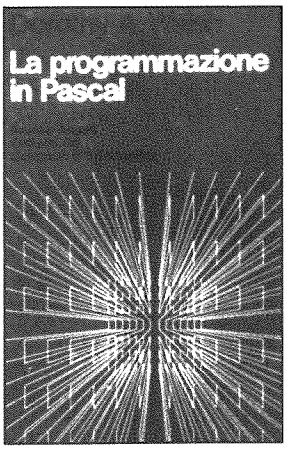


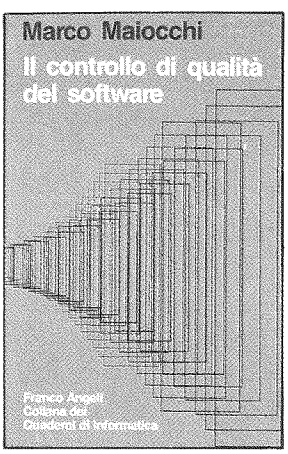


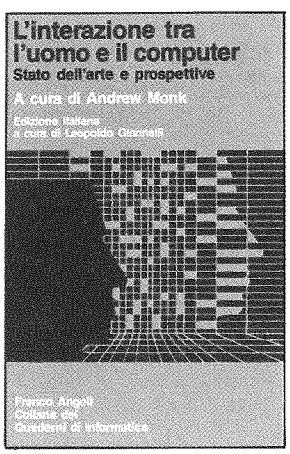


I TESTI DEI
QUADERNI DI INFORMATICA









Spedizione in abbonamento postale, Gruppo IV, 1° semestre.
Pubblicità Inferiore al 70%.
N. 35, Gennaio - Aprile 1989
TAXE PERCUE